

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: February 21, 2003, 07:40:25 ; Search time 102 Seconds
(without alignments)
1371.061 Million cell updates/sec

Title: US-09-936-145-1
Perfect score: 249
Sequence: 1 gccccgcacatacgaaga.....aaatgagaggagagatcc 249

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 442118 seqs, 280819700 residues
Total number of hits satisfying chosen parameters: 884236

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	242.8	97.5	2084	9	US-09-918-543-9
2	242.8	97.5	2084	10	US-09-769-864-58
3	242.8	97.5	2084	10	US-09-854-346-9
4	107	43.0	6837	9	US-09-928-847B-49
5	36	14.5	383	10	US-09-969-347-299
6	34.2	13.7	309	10	US-09-974-300-1061
7	34	13.7	640681	10	US-09-790-368-1
8	33.8	13.6	882	9	US-09-983-802-133
9	33.8	13.6	10716	10	US-09-954-456-270
10	33.4	13.4	516	10	US-09-864-761-7942
11	33.4	13.4	640681	10	US-09-790-988-1
12	33	13.3	425	10	US-09-834-975-451
13	32.8	13.2	2405	9	US-09-764-868-1458
14	32.8	13.2	3082	9	US-09-764-868-1459
15	32.8	13.2	9224	9	US-10-108-605-254
16	32.8	13.2	18449	9	US-09-764-868-1457
17	32.8	13.2	28588	10	US-09-764-887-399
18	32.4	13.0	18692	10	US-09-764-847-1682
19	32.4	13.0	197496	9	US-09-877-177-10

c	20	32.4	13.0	513509	9	US-09-754-853A-4	Sequence 4, Appli
c	21	31.4	12.6	2000	9	US-09-938-842A-4650	Sequence 4650, Ap
c	22	31.4	12.6	55795	10	US-09-880-107-1543	Sequence 1543, Ap
c	23	31.2	12.5	953	9	US-09-938-842A-3562	Sequence 3562, Ap
c	24	31	12.4	152	10	US-09-880-107-1627	Sequence 1627, Ap
c	25	31	12.4	31124	9	US-10-060-763-12	Sequence 12, Appli
c	26	31	12.4	31124	12	US-10-063-763-12	Sequence 12, Appli
c	27	30.8	12.4	538	10	US-09-864-761-16211	Sequence 16211, A
c	28	30.8	12.4	538	10	US-09-864-761-16454	Sequence 16454, A
c	29	30.8	12.4	9121	10	US-09-070-927A-221	Sequence 221, App
c	30	30.6	12.3	235	10	US-09-954-456-124	Sequence 124, App
c	31	30.6	12.3	766	10	US-09-764-877-958	Sequence 958, App
c	32	30.6	12.3	1120	10	US-09-764-877-3902	Sequence 3902, Ap
c	33	30.6	12.3	26197	10	US-09-764-847-1965	Sequence 1965, Ap
c	34	30.6	12.3	26210	10	US-09-764-847-1966	Sequence 1966, Ap
c	35	30.4	12.2	457	10	US-09-864-761-705	Sequence 705, App
c	36	30.4	12.2	8995	10	US-09-877-935-1	Sequence 1, Appli
c	37	30.2	12.1	394	10	US-09-880-107-1090	Sequence 1090, Ap
c	38	30.2	12.1	510	10	US-09-864-761-12921	Sequence 12921, A
c	39	30.2	12.1	783	9	US-09-822-846-22	Sequence 22, Appli
c	40	30.2	12.1	2163	10	US-09-789-561-53	Sequence 53, Appl
c	41	30.2	12.1	2309	10	US-09-789-561-79	Sequence 79, Appl
c	42	30.2	12.1	3810	10	US-09-925-301-164	Sequence 164, App
c	43	30	12.0	383	9	US-09-933-797-82	Sequence 82, Appli
c	44	30	12.0	1423	10	US-09-764-864-358	Sequence 358, App
c	45	30	12.0	2109	12	US-10-044-090-287	Sequence 287, App

ALIGNMENTS

RESULT 1
US-09-918-543-9
; Sequence 9, Application US/09918543
; Patent No. US20020155574A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020155574A1ozymes A/S
; APPLICANT: Thisted, Thomas
; APPLICANT: Kjaerulf, Soren
; APPLICANT: Andersen, Carsten
; APPLICANT: Fuglsang, Claus Crone
; TITLE OF INVENTION: Alpha-amylase mutants with altered properties
; FILE REFERENCE: 10062.200-US
; CURRENT APPLICATION NUMBER: US/09/918.543
; CURRENT FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: {343}..(1794)
; OTHER INFORMATION:
US-09-918-543-9

Query Match	97.5%	Score	242.8	DB	9	Length	2084
Best Local Similarity	99.2%	Pred.	No. 3.8e-53				
Matches	244	Conservative	0	Mismatches	2	Indels	0
Gaps	0						
QY	1	CCCCCGCATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG	60				
Db	1	CCCCCGCATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG	60				
QY	61	CTGAAGAGTGCATCGATTGTTGAGAAAGAACAGACCATATAAACCTTGTCTGTC	120				
Db	61	CTGAAGAGTGCATCGATTGTTGAGAAAGAACAGACCATATAAACCTTGTCTGTC	120				
QY	121	ATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAGAAATAGGAATAA	180				
Db	121	ATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAGAAATAGGAATAA	180				

QY 181 GGGGGTTGTTATTATTTACTGATATGATAAATAATATTTGTATATAAGAAAATGAGAGG 240
Db 181 GGGGGTTGTTATTATTTACTGATATGATAAATAATATTTGTATATAAGAAAATGAGAGG 240
QY 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 2

US-09-769-864-58
; Sequence 58, Application US/09769864
; Patent No. US20010039253A1
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben V.
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjarne
; APPLICANT: Nissen, Torben L.
; APPLICANT: Kjaerulff, Soren
; TITLE OF INVENTION: Alpha-Amulase Mutants
; FILE REFERENCE: 5368.200-US
; CURRENT APPLICATION NUMBER: US/09769,864
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/183,412
; PRIOR FILING DATE: 1998-10-30
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)...(1794)
US-09-769-864-58

Query Match 97.5%; Score 242.8; DB 10; Length 2084;
Best Local Similarity 99.2%; Pred. No. 3.8e-53;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGCCGCACATACGAAAAGACTGGCTGAAAACATTGAGCCCTTTGATGACTGATGATTGG 60
Db 1 GCGCCGCACATACGAAAAGACTGGCTGAAAACATTGAGCCCTTTGATGACTGATGATTGG 60
QY 61 CTGAGAAGTGGATCGATTGTTTCAGAAAAGAACAGACCATAAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTTCAGAAAAGAACAGACCATAAAATACCTTGTCTGTC 120
QY 121 ATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCGCTGTGFAAAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCGCTGTGFAAAAAATAGGAATAAA 180
QY 181 GGGGGTTGTTATTATTTACTGATATGATAAATAATATTTGTATATAAGAAAATGAGAGG 240
Db 181 GGGGGTTGTTATTATTTACTGATATGATAAATAATATTTGTATATAAGAAAATGAGAGG 240
QY 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 3

US-09-854-346-9
; Sequence 9, Application US/09854346
; Patent No. US20020068352A1
; GENERAL INFORMATION:
; APPLICANT: No. US20020068352A1ozymes A/S
; APPLICANT: Svendsen, Allan
; APPLICANT: Jorgensen, Christel Thea
; APPLICANT: Nielsen, Bjarne Ronfeldt
; TITLE OF INVENTION: Alpha-amylase variants with altered 1,6 activity
; FILE REFERENCE: 6140.200-US

; CURRENT APPLICATION NUMBER: US/09/854,346
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)..(1794)
; OTHER INFORMATION: BAN
US-09-854-346-9

Query Match 97.5%; Score 242.8; DB 10; Length 2084;
Best Local Similarity 99.2%; Pred. No. 3.8e-53;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCGCCGCACATACGAAAAGACTGGCTGAAAACATTGAGCCCTTTGATGACTGATGATTGG 60
Db 1 GCGCCGCACATACGAAAAGACTGGCTGAAAACATTGAGCCCTTTGATGACTGATGATTGG 60
QY 61 CTGAAGAAGTGGATCGATTGTTTCAGAAAAGAACAGACCATAAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTTCAGAAAAGAACAGACCATAAAATACCTTGTCTGTC 120
QY 121 ATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCGCTGTGFAAAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCGCTGTGFAAAAAATAGGAATAAA 180
QY 181 GGGGGTTGTTATTATTTACTGATATGATAAATAATATTTGTATATAAGAAAATGAGAGG 240
Db 181 GGGGGTTGTTATTATTTACTGATATGATAAATAATATTTGTATATAAGAAAATGAGAGG 240
QY 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 4

US-09-928-847B-49
; Sequence 49, Application US/09928847B
; Publication No. US20030032186A1
; GENERAL INFORMATION:
; APPLICANT: No. US20030032186A1ozymes A/S
; APPLICANT: Jorgensen, Steen T
; APPLICANT: Rasmussen, Michael D
; APPLICANT: Andersen, Jens Tonne
; APPLICANT: Olsen, Carsten
; TITLE OF INVENTION: Multiple Insertion of Genes
; FILE REFERENCE: 10022.204-US
; CURRENT APPLICATION NUMBER: US/09/928,847B
; CURRENT FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 49
; LENGTH: 6837
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid pMOL 1642
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (669)..(669)
; OTHER INFORMATION: n denotes an undetermined nucleotide
US-09-928-847B-49

Query Match 43.0%; Score 107; DB 9; Length 6837;
Best Local Similarity 100.0%; Pred. No. 2.8e-18;
Matches 107; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 ATGCTCTCCAGACTGCCGCTGTGTAATAAATAGGAATAAAGGGGGTGTATTATTTT 199
|||||

```
Db 382 ATGCTGTCAGACTGCTCCGCTGTGTAATAAATAGGAATAAGGGGGTGTATTATTATTTT 441
QY 200 ACTGATATGTAATAATATTTGTATAAGAAATAGAGGGGAGAGGA 246
Db 442 ACTGATATGTAATAATATTTGTATAAGAAATAGAGGGGAGAGGA 488

RESULT 5
US-09-969-347-299/c
; Sequence 299, Application US/09969347
; Patent No. US20020115085A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: Sets
; CURRENT APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/60/237,598
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,604
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 318
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 299
; LENGTH: 383
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; FEATURE:
; OTHER INFORMATION: n=a,t,g or c
US-09-969-347-299

Query Match 14.5%; Score 36; DB 10; Length 383;
Best Local Similarity 57.8%; Pred. No. 1.8;
Matches 63; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 116 CTGTCATCAGACAGGGTATTTTATGTGTCAGACTGTCGCTGTGTAAATAATAGGA 175
Db 140 CTATCAATAGTCAGTGTATTTCTTCGCGGTGCAGAAAAGNACACCAAAAAGAAAGCA 81

QY 176 ATAAAGGGGGTGTATATATTTTACTGATATGTAATAATATTTGTA 224
Db 80 ACAGGGGAGATGTTATTTATTTTAAATGACACCAAAAGTACCAATCCTA 32

RESULT 6
US-09-974-300-1061
; Sequence 1061, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: Expression
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1061
; LENGTH: 309
; TYPE: DNA
; ORGANISM: Bacillus licheniformis
US-09-974-300-1061

Query Match 13.7%; Score 34.2; DB 10; Length 309;
Best Local Similarity 60.0%; Pred. No. 5;
Matches 57; Conservative 0; Mismatches 38; Indels 0; Gaps 0;
```

```
QY 1 GCCCGCCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60
Db 214 GCCCGCCACACGATCAGCTGGCGGACGATCGATGATTTGATCGATGATGAGCTCG 273

QY 61 CTGAAGAAGTGGATCGATTTGTTTGAGAAAAGA 95
Db 274 TCCAAAGAGATCGAGCGCTTTTCCGAAAAAGA 308

RESULT 7
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. US20020127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match 13.7%; Score 34; DB 10; Length 640681;
Best Local Similarity 59.2%; Pred. No. 44;
Matches 58; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 83 TGAGAAAAGAAGACCATATAAATACCTGTCGTCTCATCAGACAGGATTTTTTATG 142
Db 156491 TGAAAAAATTTAACACCCGAAAAAATTTTGTGTCTAAGTTTTATATATTTATG 156550

QY 143 CTGTCAGACTGTCGCTGTGTAAAAATAGGAATAA 180
Db 156551 ATCAAGCGTTTACTGGATATATCGAGAATAAGAAATAA 156588
```

```
RESULT 8
US-09-983-802-133
; Sequence 133, Application US/09983802
; Publication No. US20030022185A1
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/983,802
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: EARLIER FILING DATE: 1999-01-08
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US98/13684
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,926
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,793
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,925
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,929
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,803
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,732
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,931
```

	Query Match	13.6%	Score 33.8;	DB 9;	Length 982;
	Best Local Similarity	56.9%;	Pred. No. 8.4;		
	Matches	62;	Conservative	0;	Mismatches 47; Indels 0; Gaps 0;
QY	132	TATTTTTTATGCTGCCAGACTGTCGCTGTGTAAAAAATAGGAATAAGGGGGGTGTT	191		
Db	519	TATTTTAAATCTGTTTGGATAGTCAGAAGTAGAATCATTAAGGTAAATATGAGTGT	578		
QY	192	ATTATTTTACTGATGTAAAAATAATATTTGTTATAGAAAAATGAGAGG	240		

Db 579 ACTTTGTTCTTCGATGTCATATATTTTATGTGTATATATATATGTTAAAGGG 627

RESULT 9

US-09-954-456-270

; Sequence 270, Application US/09954456

; Patent No. US20020115057A1

; GENERAL INFORMATION:

; APPLICANT: Young, Paul

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Ther

; TITLE OF INVENTION: Sets

; FILE REFERENCE: 689290-76

; CURRENT APPLICATION NUMBER: US/09/954,456

; CURRENT FILING DATE: 2001-09-18

; PRIOR APPLICATION NUMBER: US/60/233,617

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,052

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,923

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,134

; PRIOR FILING DATE: 2000-09-25

; PRIOR APPLICATION NUMBER: US/60/235,637

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,638

; PRIOR FILING DATE: 2000-09-26

; PRIOR APPLICATION NUMBER: US/60/235,711

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,720

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,840

; PRIOR FILING DATE: 2000-09-27

; PRIOR APPLICATION NUMBER: US/60/235,863

; PRIOR FILING DATE: 2000-09-27

; NUMBER OF SEQ ID NOS: 2276

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 270

; LENGTH: 10716

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-954-456-270

Query Match 13.6%; Score 33.8; DB 10; Length 1

Best Local Similarity 56.9%; Pred. No. 17;

Matches 62; Conservative 0; Mismatches 47; Indels

QY 132 TATTTTATGCTGTCACAGACTGCGCTGTGTAAAAAATAGGAATAAAGGG

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 10410 TATTTTAAATCTGTTTGATAGTACAGATAGTAATCATAAAGGTAAATATG

QY 192 ATTTTACTGATCATGTAAATATATATTTGTATAGAAAATGAGAGG 240

||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 10470 ACTTGTGTTCTTCGATGTCATATTTTATGTGTATATATATGTTAAAGG 105

RESULT 10

US-09-864-761-7942

; Sequence 7942, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC

; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

; FILE REFERENCE: Aecmica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761

; CURRENT FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 2000-05-26

RESULT 11
US-09-730-988-1/c
; Sequence 1, Application US/09790988
; Patent No. US200127687A1
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI

```

RESULT 12
US-09-834-975-451
; Sequence 451, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Brown, Jeffrey
; APPLICANT: Bolt, Andrew
; APPLICANT: Van Huffel, Christophe
; TITLE OF INVENTION: NOVEL GENES, COM
; TITLE OF INVENTION: FOR THE IDENTIF
; TITLE OF INVENTION: OF HUMAN CANCER
; FILE REFERENCE: MRI-016B
; CURRENT APPLICATION NUMBER: US/09/83
; CURRENT FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/197,538
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 1046
; SOFTWARE: FastSeq for Windows Versio
; SEQ ID NO 451
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(425)
; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-451

```

Query Match	13.3%	Score 33;	DB 10;	Length 425;
Best Local Similarity	50.7%;	Pred. No. 11;		
Matches	75; Conservative	0; Mismatches	73; Indels	0; Gaps
QY	41 TTTGTATGACTGATGATTGGCTGAAGAAGTGGATCGATTGTTTCAGAAAAGAAAGAACACC	100		
Dd	90 TTTTTTTTTTTTTTTTTTTTTNNNNAAAAAATAATTTTAAAAAATAAAAAAAAAAAAAA	149		

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 07:11:11 ; Search time 1483 Seconds
(without alignments)
2719.269 Million cell updates/sec

Title: US-09-936-145-1

Perfect score: 249

Sequence: 1 gccccgacacatcacaaaga.....aaatgagaggagagatcc 249

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: gb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pln:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	42.2	16.9	668	17 BH362752	BH362752 CH230-5L1
C 2	41.2	16.5	549	10 BE579975	BE579975 kg35c02.y
C 3	41.2	16.5	846	17 AZ684739	AZ684739 ENTMP89TR
C 4	41.2	16.5	924	17 BH158622	BH158622 ENT5V3TF
5	41	16.5	486	17 AQ266710	AQ266710 RPC111-69
6	41	16.5	575	10 BE580320	BE580320 kg39h02.y

RESULT 1	BH362752/c	BH362752	668 bp	DNA	linear	GSS 03-DEC-2001
LOCUS	CH230-5L12.TV	CHORI-230 Segment 1				Rattus norvegicus genomic clone
DEFINITION	CH230-5L12, DNA sequence.					
ACCESSION	BH362752					
VERSION	BH362752.1	GI:17293486				
KEYWORDS	GSS.					
SOURCE	Norway rat.					
ORGANISM	Rattus norvegicus					
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;					
	Rattus.					
REFERENCE	1 (bases 1 to 668)					
AUTHORS	Zhao,S., Shetty,J., Shatsman,S., Tsagaye,G., Geer,K., Shvartsbeyn					
	,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., Riggs,F., de					
	Jong,P. and Fraser,C.M.					
TITLE	Rat BAC End Sequences from Library CHORI-230 EcoRI segment					
JOURNAL	Unpublished (1999)					
COMMENT	Other_GSSs: CH230-5L12.TJ					
	Contact: Shaying Zhao					
	Department of Eukaryotic Genomics					
	The Institute for Genomic Research					
	9712 Medical Center Dr., Rockville, MD 20850, USA					
	Tel: 301 838 0200					
	Fax: 301 838 0208					
	Email: szhao@tigr.org					
	Clones are derived from the rat BAC library CHORI-230					
	(http://www.chori.org/bacpac/rat230.htm). For BAC library					

/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOS1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."
284 a 142 c 108 g 312 t

Query Match 16.5%; Score 41.2; DB 17; Length 846;
Best Local Similarity 52.3%; Pred. No. 4.1;
Matches 91; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
Qy 26 TCAAAACATTGAGCCTTTGATGACCTGATGATTTGGCTCAAGAGTGGATCGATTTTGA 85
Db 717 TAAATAATTCAACCATCTCTAATAGATGAAATGCTGAAATATAGAAAGATTTATTAA 658
Qy 86 GAAAAAGAGACACCAATAAATACCTTGTCTCATCAGACAGGCTATTTTATGCTG 145
Db 657 TAACCAAGAAACCAACAGATATTTGTTTATCTACTGACATCTTTGTTTCTGGGTG 598
Qy 146 TCCAGACTGTCGGTGTCTGTAATAAGTAAGTAAGAGGGGGTGTATTTATTTT 199
Db 597 TTCAATTTTACAAAGGATTAAAGAAAGAGGAAGTCTATTATTACTGGATT 544

RESULT 4
LOCUS BH158622/c 924 bp DNA linear GSS 24-SEP-2001
DEFINITION ENT5V33TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
genomic, DNA sequence.
ACCESSION BH158622
VERSION BH158622.1 GI:15732060
KEYWORDS GSS.
SOURCE Entamoeba histolytica.
ORGANISM Entamoeba histolytica.
REFERENCE 1 (bases 1 to 924)
AUTHORS Entamoebidae; Entamoeba.
TITLE Loftus, B., Wang, Z., Van Aken, S. and Fraser, C.
JOURNAL Determination of clone end sequences from Entamoeba histolytica
COMMENT HMI:IMSS sheared DNA library (2001)
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: bjloftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 13
High quality sequence stop: 851.

FEATURES
source
Location/Qualifiers
1..924
/organism="Entamoeba histolytica"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/note="Vector: pHOS1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.

Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for
whole genome shotgun sequencing projects. In Genome
Sequencing: A Practical Approach, eds. M. Vaudin and B.
Barell, Oxford University Press, 1999)."
318 a 152 c 112 g 342 t

Query Match 16.5%; Score 41.2; DB 17; Length 924;
Best Local Similarity 52.3%; Pred. No. 4.1;
Matches 91; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
Qy 26 TCAAAACATTGAGCCTTTGATGACCTGATGATTTGGCTCAAGAGTGGATCGATTTTGA 85
Db 483 TAAATAATTCAACCATCTCTAATAGATGAAATGCTGAAATATAGAAAGATTTATTAA 424
Qy 86 GAAAAAGAGACACCAATAAATACCTTGTCTCATCAGACAGGCTATTTTATGCTG 145
Db 423 TAACCAAGAAACCAACAGATATTTGTTTATCTACTGACATCTTTGTTTCTGGGTG 364
Qy 146 TCCAGACTGTCGGTGTCTGTAATAAGTAAGTAAGAGGGGGTGTATTTATTTT 199
Db 363 TTCAATTTTACAAAGGATTAAAGAAAGAGGAAGTCTATTATTACTGGATT 310

RESULT 5
LOCUS AQ266710 486 bp DNA linear GSS 27-APR-1999
DEFINITION RC1111-69D2.TJ RPI-11 Homo sapiens genomic clone RPI-11-69D2, DNA
sequence.
ACCESSION AQ266710
VERSION AQ266710.1 GI:3794314
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 486)
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Other_GSSs: RPI111-69D2.TK
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org

Clones are derived from the human BAC library RPI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: Sp6
Class: BAC ends.

FEATURES
source
Location/Qualifiers
1..486
/organism="Homo sapiens"
/db_xref="GDB:7526185"
/db_xref="taxon:9606"
/clone="RPI-11-69D2"
/clone_lib="RPI-11"
/sex="Male"
/cell_type="Lymphocytes"

```
/note="vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC111 Human Male BAC Library"
BASE COUNT      139 a   67 c   73 g   207 t
ORIGIN
Query Match      16.5%; Score 41; DB 17; Length 486;
Best Local Similarity 49.8%; Pred. No. 4.8;
Matches 104; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
QY 40 CTTTGATGACGTGATGTTGCTGCTGAAGAAGTGGATCGATGTTTGAGAAAGAGAGAC 99
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 16 CTTTGATTTCTTGAATTTGAATCCAAACGTTTACATCTGTATTTGTAATAAATAGGT 75
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 100 CATAAATACCTTGTCTGTCATCAGACAGGGTATTTTATGCTGCCAGACTGCCGC 159
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 76 TATATAATTTTCCCTATTTTGTACTGCTTGTGAATTTTATATTAAGGCTATAGA 135
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 160 TGTGTAATAATAGATAAAGGGGGTGTCTATATTTTACTGATATGTAATAATAAT 219
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 136 TCTCAGCAATAAATACATCTCTTTATGTTTATATTCCTTTCTTAACACAATAAAT 195
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 220 TTGCTATAAGAAATCAGAGGGGAGAGGATC 248
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 196 TTATTCATTATGTTGCATGCCAGTAGTTC 224
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||

RESULT 6
BE580320
LOCUS      BE580320          575 bp      mRNA      linear      EST 09-MAY-2001
DEFINITION      kq39h02.y1 TBN95TM-SSR Strongyloides stercoralis cDNA 5' similar to
                  WP:2C247.1 CE19306 ;, mRNA sequence.
ACCESSION      BE580320
VERSION        BE580320.1 GI:9831262
KEYWORDS       EST.
SOURCE         Strongyloides stercoralis.
ORGANISM       Strongyloides stercoralis
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida;
Panagrolaimoidea; Strongyloidea; Strongyloidea.
REFERENCE      1 (bases 1 to 575)
AUTHORS        McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,
                  Dante,M., Marra,M., Hillier,L., Kucaba,I., Theising,B., Bowers,Y.,
                  Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagaris,I., R.,
                  Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
                  ,M., Allen,M., Persson,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
                  Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
                  Wilson,R.
                  The Washington Univ. Nematode EST Project, 1999
                  Unpublished (1999)
                  Contact: McCarter JP
                  The Washington Univ. Nematode EST Project, 1999
                  Washington University School of Medicine
                  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                  Tel: 314 286 1800
                  Fax: 314 286 1810
                  Email: est@watson.wustl.edu
                  The library was constructed by Dr. Thomas Nutman and colleagues of
                  NIAID, NIH (tnutman@nih.gov). DNA Sequencing by: Washington
                  University Genome Sequencing Center St. Louis.
                  High quality sequence stop: 421.
FEATURES       source
                  Location/Qualifiers
                    1..575
                      /organism="Strongyloides stercoralis"
                      /strain="Rhabditiform larvae obtained from gerbils"
                      /db_xref="taxon:6248"
                      /clone_lib="TBN95TM-SSR"
                      /lab_host="XL-1 Blue MRF" (Stratagene)"
                      /note="Vector: Lambda Uni-ZAP XR (Stratagene); Site_1:
EcoRI; Site_2: XhoI; mRNA was purified from 2 x 10E3
rhabditiform larvae which had been isolated from gerbils
experimentally infected with larvae originally isolated
from experimentally infected dogs. cDNA was constructed
and, using adaptors, was cloned unidirectionally into the
vector from the EcoRI site to the XhoI site. The library
```

```
has an unamplified titer of 1 x 10E5 pfu/ml and an
amplified, undiluted titer of 9 x 10E11 pfu/ml. The
average insert size of the unamplified library is 675 bp
(range, 100-1700)."
BASE COUNT      237 a   74 c   106 g   157 t   1 others
ORIGIN
Query Match      16.5%; Score 41; DB 10; Length 575;
Best Local Similarity 55.8%; Pred. No. 4.7;
Matches 77; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 16 ARAAGACTGGCTGAAAACATTTGAGCCCTTTGATGACTGATGATTTGGCTGAAGAAGTGGATC 75
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
Db 409 ATAAACTTGTGTAAGAGGATGAACCACTTAAGACTGTTGATAATGATAACAGCAAGAAA 468
      ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||
QY 76 GATTGTTTGAGAAAAGAAGACCATATAAAATACCTTGTCTGTCATCAGACAGGGTATT 135
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 469 TATTGTTGTAANAAGAAAACACCTCGAAGATGACATTTTGGCATATATTTAAAGAGATT 528
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 136 TTTTATGCTGTCCAGACT 153
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 529 TATCAGATTATCCAAAAT 546
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 7
CNS017KX/c
CNS017KX/c
LOCUS      CNS017KX          1101 bp      DNA      linear      GSS 26-JUL-1999
DEFINITION      Drosophila melanogaster genome survey sequence SP6 end of BAC
                  BAC37F10 of DrosBAC library from Drosophila melanogaster (fruit
                  fly), genomic survey sequence.
ACCESSION      AL108171
VERSION        AL108171.1 GI:5628475
KEYWORDS       GSS.
SOURCE         Drosophila melanogaster.
ORGANISM       Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 1101)
AUTHORS        Genoscope.
TITLE          Direct Submission
JOURNAL        Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
                  BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
                  - Web : www.genoscope.cns.fr)
COMMENT        Determination of this BAC-end sequence was carried out as part of a
                  collaboration with the European Drosophila Genome Project (EDGP) -
                  http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
                  library (Dros BAC) was made by Alain Billaud at CEPH (Centre
                  d'Etude du Polymorphisme Humain) with funding provided by a MRC
                  project grant. The DNA was prepared from embryos by Alain Bucheton
                  and Genevieve Payan. It has been constructed in the vector
                  pBelobAC11.
FEATURES       Location/Qualifiers
                  source
                    1..1101
                      /organism="Drosophila melanogaster"
                      /db_xref="taxon:7227"
                      /clone="BACN37F10"
                      /clone_lib="DrosBAC"
                      /plasmid="pBelobAC11"
                      /note="end : SP6"

BASE COUNT      247 a   84 c   159 g   263 t   348 others
ORIGIN
Query Match      16.0%; Score 39.8; DB 17; Length 1101;
Best Local Similarity 18.4%; Pred. No. 9;
Matches 41; Conservative 107; Mismatches 74; Indels 1; Gaps 1;
QY 22 TGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGGCTGAAGAAGTGGATCGATTGT 81
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1099 KDDDKDDTDRFTWTKRKKKTTWKTDKKWRWDTKDGGGGGGKKGKGGDRADTK 1040
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
QY 82 TTGAGAAAAGAGAACCAATAAATACCTTGTCTGTCATCAGACAGGGTATTTTAT 141
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
```

```

Db 1039 RDWRGAATRRARWATAAATATADDTWKRTTGKTKTKWKKWWARAAKWKDWRDNDWD 980

Qy 142 G-CTCCAGACTGTCGCTGTGTAATAAATAGGAATAAGGGGGTGTGTTATTATTTTA 200
: |:: : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Db 979 KWTWKWDRTWKTRATRGDGRAGRAKRRDRGRGRRRRRTTKRWWRRAAWTTAW 920
: |:: : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Qy 201 CTGATATGTAAATATTAATTTGTTATAGAAAATAGAGGGGAGA 243
: |:: : : : : | : : : : | : : : : | : : : : | : : : : | : : : :
Db 919 TTTWWGTTTGAATTKRTRKGGGAAADWGTDDWRWDKRAA 877

RESULT 8
AW292419 392 bp mRNA linear EST 16-JAN-2000
LOCUS UI-H-BI2-agw-h-08-0-UI.s1 NCI_CGAP_Sub4 Homo sapiens cDNA clone
DEFINITION IMAGE:2725886 3', mRNA sequence.
ACCESSION AW292419
VERSION AW292419.1 GI:6699055
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 392)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution:
NCI-CGAP clone distribution information can be found through the
I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html The following repetitive
elements were found in this cDNA sequence: 1-30,
>AT_rich#Low_complexity
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
Location/Qualifiers
1..392
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2725886"
/clone_lib="NCI_CGAP_Sub4"
/lab_host="DH10B (Life Technologies)"
/notes="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; The
NCI_CGAP_Sub4 library is a subtracted library derived from
the NCI_CGAP_Sub2 library which is a subtracted library
derived from the NCI_CGAP_Sub1 library, which is a
subtracted library derived from BI. BI constitutes a
mixture of 21 normalized or subtracted NCI_CGAP
libraries: NCI_CGAP_Co4, NCI_CGAP_Pr22, NCI_CGAP_Pr28,
NCI_CGAP_Co10, NCI_CGAP_Co16, NCI_CGAP_Kid5,
NCI_CGAP_Kid12, NCI_CGAP_Kid3, NCI_CGAP_Kid11,
NCI_CGAP_Lym2, NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_CLL1,
NCI_CGAP_Leiz2, NCI_CGAP_Brn23, NCI_CGAP_Lu5,
NCI_CGAP_Lu24, NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6,
NCI_CGAP_Brn25. These 21 libraries were pooled and a
single-stranded DNA preparation of the resulting mixture
was used as a tracer in a subtractive hybridization with
a driver whose composition is detailed below:
NCI_CGAP_Kid3 pool 1 : LLAM 3334-3337, 3682-3683,
3798-3803 (IMAGE Clones 1322376-1323911,
1456008-1456775, 1500552-1502855) NCI_CGAP_Kid5 pool 1 :
LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clones
1323912-1325831, 1471368-1472903, 1492104-1493255)
NCI_CGAP_Lu5 pool 1 : LLAM 3575-3582, 3851-3854 (IMAGE
Clones 1414920-1417991, 1520904-1522439) NCI_CGAP_GC4
pool 1 : LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE

```

```

CloneIDs 1257096-1258631, 1469064-1470983, 1475592-1476743
) NCI_CGAP_Pr22 pool 1 : LLAM 2457-2459, 2758-2759,
3062-3068 (IMAGE Clones 985608-986759, 1101192-1101959,
1217928-1220615) NCI_CGAP_Co10 pool 1 : LLAM 2644-2653,
2871-2872 (IMAGE Clones 1057416-1061255, 114584-1145351
) Subtraction was performed as previously described
(Bonaldo, Lennon & Soares (1996)): Normalization and
Subtraction: Two Approaches To Facilitate Gene Discovery.
Genome Research 6, 791-806.]
TAG LIB=NCI_CGAP_Lu5
TAG_TISSUE=lung
TAG_SEQ=CAAC"

BASE COUNT 107 a 113 g 91 t
ORIGIN
Query Match 15.7%; Score 39.2; DB 10; Length 392;
Best Local Similarity 73.5%; Pred. No. 13;
Matches 50; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 166 AAAATAGGAATAAAGGGGGTGTATTATTTACTGATGTAATAATTTGAT 225
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 193 AAAGAGAGGAGAACTAAGGTGTCTTAATGACAGATAAGTAAATAATTCCTT 252
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 226 AAGAAAT 233
| | | | | | | |
Db 253 AAGAAAT 260

RESULT 9
AW001963 461 bp mRNA linear EST 09-MAR-2000
LOCUS wt82d01.x1 Soares_thymus_NHFTh Homo sapiens cDNA clone
DEFINITION IMAGE:2513953 3', mRNA sequence.
ACCESSION AW001963
VERSION AW001963.1 GI:5848879
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 461)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 696 Std Error: 0.00
Seq primer: -40Up from Gibco
High quality sequence stop: 448.
Location/Qualifiers
1..461
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2513953"
/clone_lib="Soares_thymus_NHFTh"
/dev_stage="fetal"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: thymus, pooled; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCAACGTTTTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 141 a 107 c 127 g 86 t
ORIGIN

```

Matches	50;	Conservative		0;	Mismatches	18;	Indels	0;	Gaps	0;
QY	166	AAAAATAGGAATAAAGGGGTTGTATTACTTATCTATGTAATAAAATAATTTGTAT	225 							
Dd	270	AAGAAGAGGAGAACTAAGTGTCATTAAATACAGATAAGTAAAATAATATCGCTTT	211 							
QY	226	AGAAANAAT 233								
Dd	210	AAGAAANAT 203								
RESULT 11										
AI989386										
LOCUS wt80e12.x1 Soares.thymus_NHfTh Homo sapiens CDNA clone EST 08-SEP-1999										
DEFINITION IMAGE:2513806 3' , mRNA sequence.										
ACCESSION AI989386										
VERSION AI989386.1 GI:5836309										
KEYWORDS EST.										
SOURCE human.										
ORGANISM Homo sapiens										
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 467)										
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.										
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), TITLE Tumor Gene Index										
JOURNAL Unpublished (1997)										
COMMENT Contact: Robert Strauszberg, Ph.D. Email: cgaaps.r@mail.nih.gov This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: -40UP from Gibco High quality sequence stop: 442. Location/Qualifiers I . .467 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2513806" /clone_lib="Soares_thymus_NHFTh" /dev_stage="fetal" /lab_host="DH10B (phage-resistant)" /note="Organ: thymus, pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCATTCTCAGTAGGCGGCCGCACAGTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo." BASE COUNT 144 a 110 c 127 g 86 t ORIGIN Query Match 15.7%; Score 39.2; DB 9; Length 467; Best Local Similarity 73.5%; Pred. No. 13; Matches 50; Conservative 0; Mismatches 18; Indels 0; Gaps 0; QY 166 AAAAAATAGGAATAAAGGGGTTGTATTACTTATCTATGTAATAAAATAATTTGTAT 225 Dd 176 AAGAAGAGGAGAACTAAGTGTCATTAAATACAGATAAGTAAAATAATATCGCTTT 235 QY 226 AGCAANAAT 233 Dd 236 AAGAAANAT 243 RESULT 12 AI97642 LOCUS we15g02.x1 NCI_CGAP_Lu24 Homo sapiens cDNA clone IMAGE:2341202 3' , DEFINITION mRNA sequence.										

```

ACCESSION   AI697642
VERSION     AI697642.1  GI:4985542
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 475)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            DNA Library Arrayed by: Greg Lennon, Ph.D.
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 701 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 465.
FEATURES    Location/Qualifiers
            1..475
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:2341202"
            /clone_lib="NCI-CGAP_Lu24"
            /tissue_type="carcinoid"
            /lab_host="DH10B"
            /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
            modified polylinker; Plasmid DNA from the normalized
            library NCI-CGAP_Lu5 was prepared and ss circles were
            made in vitro. Following HAP purification, this DNA was
            used as tracer in a subtractive hybridization reaction.
            The driver was PCR-amplified cDNAs from a pool of 5,000
            clones made from the same library (cloneIDs
            1414920-1417991 and 1520904-1522439). Subtraction by Bento
            Soares and M. Fatima Bonaldo."
BASE COUNT  147 a 109 c 130 g 89 t
ORIGIN
Query Match      15.7%; Score 39.2; DB 9; Length 475;
Best Local Similarity 73.5%; Pred. No. 13;
Matches 50; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 166 AAAATAGGAATAAGGGGGTGTATTATTTACTGATATGTAATAATAATTGTAT 225
    ||| |||| ||| ||||| |||| | |||| ||||| ||||| |||
Db 179 AAAGAGAGGAGAACTAAGGTGTTCATTAAATGACAGATAAGTAAAAATAATTGCTT 238
    ||| |||| ||| ||||| |||| | |||| ||||| ||||| |||

Qy 226 AAGAAAT 233
    |||||
Db 239 AAGAAAT 246

RESULT 13
AI479617
LOCUS       tm32b12.x1 NCI-CGAP_CLL1 Homo sapiens cDNA clone IMAGE:2158271 3',
DEFINITION mRNA sequence.
ACCESSION   AI479617
VERSION     AI479617.1  GI:4372785
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 564)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            DNA Library Arrayed by: Greg Lennon, Ph.D.
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 701 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 465.
FEATURES    Location/Qualifiers
            1..564
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:2158271"
            /clone_lib="NCI-CGAP_CLL1"
            /tissue_type="B-cell, chronic lymphocytic leukemia"
            /lab_host="DH10B"
            /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
            polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
            was primed with a Not I - oligo(dT) primer [5',
            TGTTACCAATCTGAAGTGGAGCGCGCATGTGTTTTTTTTTTTTTTT
            T 3']; double-stranded cDNA was ligated to Eco RI
            adaptors (Pharmacia), digested with Not I and cloned into
            the Not I and Eco RI sites of the modified pT7T3 vector.
            Library is normalized, and was constructed by Bento
            Soares and M. Fatima Bonaldo."
BASE COUNT  177 a 133 c 147 g 107 t
ORIGIN
Query Match      15.7%; Score 39.2; DB 9; Length 564;
Best Local Similarity 73.5%; Pred. No. 13;
Matches 50; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 166 AAAATAGGAATAAGGGGGTGTATTATTTACTGATATGTAATAATAATTGTAT 225
    ||| |||| ||| ||||| |||| | |||| ||||| ||||| |||
Db 173 AAAGAGAGGAGAACTAAGGTGTTCATTAAATGACAGATAAGTAAAAATAATTGCTT 232
    ||| |||| ||| ||||| |||| | |||| ||||| ||||| |||

Qy 226 AAGAAAT 233
    |||||
Db 233 AAGAAAT 240

RESULT 14
CNS0182P
LOCUS       BACN37D10 of DrosBAC library from Drosophila melanogaster (fruit
DEFINITION fly), genomic survey sequence.
ACCESSION   ALI08811
VERSION     ALI08811.1  GI:5629115
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
ORGANISM    Drosophila melanogaster.
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL     Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
            Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC

```


GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 07:16:20 ; Search time 76 seconds
(without alignments)
1004.770 Million cell updates/sec

Title: US-09-936-145-1

Perfect score: 249

Sequence: 1 gccccgcacatacgaaga.....aatgagaggagagatcc 249

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PCRUS_COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	246	98.8	2604	4	US-09-264-097-3
2	246	98.8	2604	4	US-09-537-168-5
3	242.8	97.5	2084	1	US-08-720-899-3
4	242.8	97.5	2084	1	US-08-459-610-3
5	242.8	97.5	2084	2	US-08-343-804-3
6	242.8	97.5	2084	2	US-08-687-399-3
7	242.8	97.5	2084	2	US-08-600-908A-3
8	242.8	97.5	2084	3	US-08-683-838A-3
9	242.8	97.5	2084	3	US-09-182-859-3
10	242.8	97.5	2084	4	US-09-170-670-13
11	242.8	97.5	2084	4	US-09-193-068-31
12	242.8	97.5	2084	4	US-09-183-412-58
13	242.8	97.5	2084	4	US-09-290-734-13
14	242.8	97.5	2084	4	US-09-672-459-3
15	242.8	97.5	2084	4	US-09-636-252A-3
16	170.8	68.6	10216	2	US-08-875-154-1
17	162	65.1	185	4	US-09-384-305-20
18	162	65.1	185	4	US-09-258-377-25
19	158.8	63.8	162	1	US-08-434-255-24
20	158.8	63.8	162	1	US-08-459-967-24
21	158.8	63.8	162	1	US-08-460-327-24
22	158.8	63.8	162	1	US-08-459-871-24
23	157.2	63.1	185	4	US-09-384-305-22
24	157.2	63.1	185	4	US-09-258-377-27
25	155.6	62.5	185	4	US-09-384-305-21
26	155.6	62.5	185	4	US-09-258-377-26
27	39.2	15.7	44	4	US-09-384-305-15

28	39.2	15.7	44	4	US-09-258-377-19	Sequence 19, Appl
29	36.4	14.6	1511	1	US-07-991-867B-8	Sequence 8, Appl
30	36.4	14.6	1511	1	US-08-107-755A-8	Sequence 8, Appl
31	36.4	14.6	1511	2	US-08-544-332-8	Sequence 8, Appl
32	36.4	14.6	1511	4	US-09-370-861A-8	Sequence 8, Appl
c 33	35.4	14.2	8700	2	US-08-392-625-16	Sequence 16, Appl
c 34	35.4	14.2	8700	2	US-08-466-961A-18	Sequence 18, Appl
c 35	35.4	14.2	8700	2	US-08-645-193B-16	Sequence 16, Appl
c 36	34	13.7	34	1	US-08-434-255-26	Sequence 26, Appl
c 37	34	13.7	34	1	US-08-459-967-26	Sequence 26, Appl
c 38	34	13.7	34	1	US-08-460-327-26	Sequence 26, Appl
c 39	34	13.7	34	1	US-08-459-871-26	Sequence 26, Appl
c 40	34	13.7	43	2	US-08-873-479-38	Sequence 38, Appl
c 41	34	13.7	43	2	US-09-031-442A-10	Sequence 10, Appl
c 42	34	13.7	43	2	US-08-972-661A-26	Sequence 26, Appl
c 43	34	13.7	43	4	US-09-258-377-10	Sequence 10, Appl
44	33.8	13.6	882	4	US-09-227-357-133	Sequence 133, App
45	33.4	13.4	9717	4	US-09-251-645-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-09-264-097-3

; Sequence 3, Application US/09264097

; Patent No. 6287826

; GENERAL INFORMATION:

; APPLICANT: No. 6287826man, Barrie Edmund

; APPLICANT: Hendriksen, Hanne Vang

; TITLE OF INVENTION: Enzymatic Preparation of Glucose Syrup

; FILE REFERENCE: 5278.200-US

; CURRENT APPLICATION NUMBER: US/09/264,097

; CURRENT FILING DATE: 1999-03-08

; EARLIER APPLICATION NUMBER: PA 0321/98

; EARLIER FILING DATE: 1998-03-09

; EARLIER APPLICATION NUMBER: 60/079,209

; EARLIER FILING DATE: 1998-03-24

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 2604

; TYPE: DNA

; ORGANISM: Bacillus amyloliquefaciens

; US-09-264-097-3

Query Match 98.8%; Score 246; DB 4; Length 2604;
Best Local Similarity 100.0%; Pred. No. 1.9e-59;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCCTTTGATGACTGATGATTGG 60

Db 521 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCCTTTGATGACTGATGATTGG 580

Qy 61 CTGAAGAAGTGGATCGATTGTTTCAGAAAACAAGAACACCAATAACCTTCTGCTGTC 120

Db 581 CTGAAGAAGTGGATCGATTGTTTCAGAAAACAAGAACACCAATAACCTTCTGCTGTC 640

Qy 121 ATCAGACAGGGTATTTTATGCTGTCGAGACTGTCGCGTGTGTAATAAATAGGAATAAA 180

Db 641 ATCAGACAGGGTATTTTATGCTGTCGAGACTGTCGCGTGTGTAATAAATAGGAATAAA 700

Qy 181 GGGGGTGTATTATTTTACTGATGTAATAATATTAATTTGTATAGAAATGAGAGGG 240

Db 701 GGGGGTGTATTATTTTACTGATGTAATAATATTAATTTGTATAGAAATGAGAGGG 760

Qy 241 AGAGGA 246

Db 761 AGAGGA 766

RESULT 2

```
US-09-537-168-5
; Sequence 5, Application US/09537168
; Patent No. 6410295
; GENERAL INFORMATION:
; APPLICANT: Andersen, Carsten
; APPLICANT: Jorgensen, Christel T.
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjaerulff, Soren
; TITLE OF INVENTION: Alpha-Amylase Variants
; FILE REFERENCE: 5886.200-US
; CURRENT APPLICATION NUMBER: US/09/537,168
; CURRENT FILING DATE: 2000-03-29
; EARLIER APPLICATION NUMBER: PA 1999 00437
; EARLIER FILING DATE: 1999-03-30
; EARLIER APPLICATION NUMBER: 60/127,427
; EARLIER FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2604
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
US-09-537-168-5

Query Match          98.8%; Score 246; DB 4; Length 2604;
Best Local Similarity 100.0%; Pred. No. 1.9e-59;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTTGG 60
Db 521 GCCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTTGG 580
Qy 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGCTGTC 120
Db 581 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGCTGTC 640
Qy 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCCGTGTGTAATAAATAGGAATAAA 180
Db 641 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCCGTGTGTAATAAATAGGAATAAA 700
Qy 181 GGGGGTGTATTTATTTACTGATGTAAATAATAATTTGTATAAGAAATAGAGAGG 240
Db 701 GGGGGTGTATTTATTTACTGATGTAAATAATAATTTGTATAAGAAATAGAGAGG 760
Qy 241 AGAGGA 246
Db 761 AGAGGA 766

RESULT 3
US-08-720-899-3
; Sequence 3, Application US/08720899
; Patent No. 5753460
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57534600 No. 57534600disk of No. 5753460th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,899
; FILING DATE: 10-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343,804
; FILING DATE: 22-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lowney Dr., Karen A.
; REGISTRATION NUMBER: 31,274
; REFERENCE/DOCKET NUMBER: 4054.214-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2084 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 250..1794
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 250..342
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 343..1791
US-08-720-899-3

Query Match          97.5%; Score 242.8; DB 1; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.3e-58;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTTGG 60
Db 1 GCCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTTGG 60
Qy 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAGACCAATAAAATACCTTGCTGTC 120
Qy 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCCGTGTGTAATAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCCGTGTGTAATAAATAGGAATAAA 180
Qy 181 GGGGGTGTATTTATTTACTCATATGTAATAATAATTTGTATAAGAAATAGAGAGG 240
Db 181 GGGGGTGTATTTATTTACTCATATGTAATAATAATTTGTATAAGAAATAGAGAGG 240
Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 4
US-08-459-610-3
; Sequence 3, Application US/08459610
; Patent No. 5801043
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58010430 No. 5801043disk of No. 5801043th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
```


RESULT 5
US-08-343-804-3
; Sequence 3, Application US/08343804
; Patent No. 5830837
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, H
; APPLICANT: Borchert, Torben Ved
; APPLICANT: Svendsen, Allan
; APPLICANT: Thøllersen, Marianne

Db 241 AGAGGA 246

RESULT 8

US-08-683-838A-3
 ; Sequence 3, Application US/08683838A
 ; Patent No. 6022724
 ; GENERAL INFORMATION:
 ; APPLICANT: Svendsen, Allan
 ; APPLICANT: Bisg rd-Frantzen, Henrik
 ; APPLICANT: Borchert, Torben Vedel
 ; TITLE OF INVENTION: "-Amylase Mutants
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: NO. 60227240 NO. 6022724disk of No. 6022724th America, Inc.
 ; STREET: 405 Lexington Avenue, 64th Floor
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: United States of America
 ; ZIP: 10174-6401
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/683,838A
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/600,908
 ; FILING DATE: 13-FEB-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Green, Reza
 ; REGISTRATION NUMBER: 38,475
 ; REFERENCE/DOCKET NUMBER: 4394.204-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212-867-0123
 ; TELEFAX: 212-878-9655
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2084 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 250..1791
 ; FEATURE:
 ; NAME/KEY: sig_peptide
 ; LOCATION: 250..342
 ; FEATURE:
 ; NAME/KEY: mat_peptide
 ; LOCATION: 343..1791
 ; US-08-683-838A-3

Query Match 97.5%; Score 242.8; DB 3; Length 2084;
 Best Local Similarity 99.2%; Pred. No. 1.3e-58;
 Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCCTTTGATGACTGATGATTGG 60
 Db 1 GCCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCCTTTGATGACTGATGATTGG 60
 Qy 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAACACCATAAAATACCTGCTGTC 120
 Db 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAACACCATAAAATACCTGCTGTC 120
 Qy 121 ATCAGACAGGGTATTTTTTATGCTGCCAGACTGCTCGCTGTGTAAGAAATAGCAATAAA 180
 Db 121 ATCAGACAGGGTATTTTTTATGCTGCCAGACTGCTCGCTGTGTAAGAAATAGCAATAAA 180

Qy 181 GGGGGTGTATTATTACTGATATGTAATAATATAATTTGTATAAGAAAATGAGAGGG 240
 Db 181 GGGGGTGTATTATTACTGATATGTAATAATATAATTTGTATAAGAAAATGAGAGGG 240
 Qy 241 AGAGGA 246
 Db 241 AGAGGA 246

RESULT 9

US-09-182-859-3
 ; Sequence 3, Application US/09182859
 ; Patent No. 6143708
 ; GENERAL INFORMATION:
 ; APPLICANT: Svendsen, Allan
 ; APPLICANT: Borchert, Torben
 ; APPLICANT: Bisgard-Frantzen, Henrik
 ; TITLE OF INVENTION: Alpha-Amylase Mutants
 ; FILE REFERENCE: 4796.204-US
 ; CURRENT APPLICATION NUMBER: US/09/182,859
 ; CURRENT FILING DATE: 1998-10-29
 ; EARLIER APPLICATION NUMBER: 0515/96
 ; EARLIER FILING DATE: 1996-04-30
 ; EARLIER APPLICATION NUMBER: 0712/96
 ; EARLIER FILING DATE: 1996-06-28
 ; EARLIER APPLICATION NUMBER: 0775/96
 ; EARLIER FILING DATE: 1996-07-11
 ; EARLIER APPLICATION NUMBER: 1263/96
 ; EARLIER FILING DATE: 1996-11-08
 ; NUMBER OF SEQ ID NOS: 37
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 3
 ; LENGTH: 2084
 ; TYPE: DNA
 ; ORGANISM: Bacillus amyloliquefaciens
 ; US-09-182-859-3

Query Match 97.5%; Score 242.8; DB 3; Length 2084;
 Best Local Similarity 99.2%; Pred. No. 1.3e-58;
 Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCCTTTGATGACTGATGATTGG 60
 Db 1 GCCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCCTTTGATGACTGATGATTGG 60
 Qy 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAACACCATAAAATACCTGCTGTC 120
 Db 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAACACCATAAAATACCTGCTGTC 120
 Qy 121 ATCAGACAGGGTATTTTTTATGCTGCCAGACTGCTCGCTGTGTAAGAAATAGCAATAAA 180
 Db 121 ATCAGACAGGGTATTTTTTATGCTGCCAGACTGCTCGCTGTGTAAGAAATAGCAATAAA 180
 Qy 181 GGGGGTGTATTATTACTGATATGTAATAATATAATTTGTATAAGAAAATGAGAGGG 240
 Db 181 GGGGGTGTATTATTACTGATATGTAATAATATAATTTGTATAAGAAAATGAGAGGG 240
 Qy 241 AGAGGA 246
 Db 241 AGAGGA 246

RESULT 10

US-09-170-670-13
 ; Sequence 13, Application US/09170670
 ; Patent No. 6187576
 ; GENERAL INFORMATION:
 ; APPLICANT: Svendsen, Allan
 ; APPLICANT: Borchert, Torben
 ; APPLICANT: Bisgard-Frantzen, Henrik
 ; TITLE OF INVENTION: Alpha-Amylase Mutants
 ; FILE REFERENCE: 5276.200-US
 ; CURRENT APPLICATION NUMBER: US/09/170,670

```
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 1172/97
; EARLIER FILING DATE: 1997-10-13
; EARLIER APPLICATION NUMBER: 60/063,306
; EARLIER FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: B. amyloliquefaciens
us-09-170-670-13

Query Match          97.5%; Score 242.8; DB 4; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.3e-58;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60

Qy 61 CTGAAGAAGTGGATCGATGTTTTCAGAAAAGAGAAGACACATAAAATACCTTGTCTGC 120
Db 61 CTGAAGAAGTGGATCGATGTTTTCAGAAAAGAGAAGACACATAAAATACCTTGTCTGC 120

Qy 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAATAGGAATAAA 180

Qy 181 GGGGGTGTGTTATTTACTGATATGTAATAATAATTTCTATAGAAAATGAGAGGG 240
Db 181 GGGGGTGTGTTATTTACTGATATGTAATAATAATTTCTATAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 11
US-09-193-068-31
; Sequence 31, Application US/09193068
; Patent No. 6197565
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjrulff, S ren
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Andersen, Carsten
; TITLE OF INVENTION: -Amylase Variants
; FILE REFERENCE: 5709.000-US
; CURRENT APPLICATION NUMBER: US/09/193,068
; CURRENT FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: B. amyloliquefaciens
US-09-193-068-31

Query Match          97.5%; Score 242.8; DB 4; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.3e-58;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60

Qy 61 CTGAAGAAGTGGATCGATGTTTTCAGAAAAGAGAAGACACATAAAATACCTTGTCTGC 120
Db 61 CTGAAGAAGTGGATCGATGTTTTCAGAAAAGAGAAGACACATAAAATACCTTGTCTGC 120

Qy 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAATAGGAATAAA 180
```

```
Db 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAATAGGAATAAA 180
Qy 181 GGGGGTGTGTTATTTACTGATATGTAATAATAATTTCTATAGAAAATGAGAGGG 240
Db 181 GGGGGTGTGTTATTTACTGATATGTAATAATAATTTCTATAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 12
US-09-183-412-58
; Sequence 58, Application US/09183412
; Patent No. 6204232
; GENERAL INFORMATION:
; APPLICANT: Borchert, Torben V.
; APPLICANT: Svendsen, Allan
; APPLICANT: Andersen, Carsten
; APPLICANT: Nielsen, Bjarne
; APPLICANT: Nissen, Torben L.
; APPLICANT: Kjaerulff, Soren
; TITLE OF INVENTION: Alpha-Amulase Mutants
; FILE REFERENCE: 5368.200-US
; CURRENT APPLICATION NUMBER: US/09/183,412
; CURRENT FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: 60/064,662
; EARLIER FILING DATE: 1997-11-06
; EARLIER APPLICATION NUMBER: 60/093,234
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 1240/97
; EARLIER FILING DATE: 1997-10-30
; EARLIER APPLICATION NUMBER: PA 1998 00936
; EARLIER FILING DATE: 1998-07-14
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 58
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (343)...(1794)
US-09-183-412-58

Query Match          97.5%; Score 242.8; DB 4; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.3e-58;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60

Qy 61 CTGAAGAAGTGGATCGATGTTTTCAGAAAAGAGAAGACACATAAAATACCTTGTCTGC 120
Db 61 CTGAAGAAGTGGATCGATGTTTTCAGAAAAGAGAAGACACATAAAATACCTTGTCTGC 120

Qy 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCGCTGTGTAATAAATAGGAATAAA 180

Qy 181 GGGGGTGTGTTATTTACTGATATGTAATAATAATTTCTATAGAAAATGAGAGGG 240
Db 181 GGGGGTGTGTTATTTACTGATATGTAATAATAATTTCTATAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 13
US-09-290-734-13
; Sequence 13, Application US/09290734
```

```
; Patent No. 6361989
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Outtrup, Helle
; APPLICANT: Nielsen, Bjarne Ronfeldt
; APPLICANT: Nielsen, Vibeke Skovgaard
; APPLICANT: Hoeck, Lisbeth Hedegaard
; TITLE OF INVENTION: No. 6361989el -Amylase And -Amylase Mutants
; FILE REFERENCE: 5276.400-US
; CURRENT APPLICATION NUMBER: US/09/290,734
; CURRENT FILING DATE: 1999-04-13
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: B. amyloliquefaciens
; US-09-290-734-13

Query Match          97.5%; Score 242.8; DB 4; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.3e-58;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG 60

Qy 61 CTGAAGAAGTGGATCGATTGTTGAGAAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTGAGAAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTAAGAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTAAGAAAATAGGAATAAA 180

Qy 181 GGGGGTGTGTTATTTACTGATATGTAAATATAATTTGTATAAGAAAATGAGAGGG 240
Db 181 GGGGGTGTGTTATTTACTGATATGTAAATATAATTTGTATAAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 14
US-09-459-3
; Sequence 3, Application US/09672459
; Patent No. 6436888
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Borchert, Torben
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 4796.204-US
; CURRENT APPLICATION NUMBER: US/09/672,459
; CURRENT FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: 09/182,859
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 0515/96
; PRIOR FILING DATE: 1996-04-30
; PRIOR APPLICATION NUMBER: 0712/96
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 0775/96
; PRIOR FILING DATE: 1996-07-11
; PRIOR APPLICATION NUMBER: 1263/96
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2084
; TYPE: DNA

Query Match          97.5%; Score 242.8; DB 4; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.3e-58;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG 60

Qy 61 CTGAAGAAGTGGATCGATTGTTGAGAAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTGAGAAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTAAGAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTAAGAAAATAGGAATAAA 180

Qy 181 GGGGGTGTGTTATTTACTGATATGTAAATATAATTTGTATAAGAAAATGAGAGGG 240
Db 181 GGGGGTGTGTTATTTACTGATATGTAAATATAATTTGTATAAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 15
US-09-636-252A-3
; Sequence 3, Application US/09636252A
; Patent No. 6440716
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 0776/1F216-US2
; CURRENT APPLICATION NUMBER: US/09/636,252A
; CURRENT FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 08/683,838
; PRIOR FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: B. amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (250)...(1791)
; NAME/KEY: mat_peptide
; LOCATION: (343)...(1791)
; NAME/KEY: sig_peptide
; LOCATION: (250)...(342)
; US-09-636-252A-3

Query Match          97.5%; Score 242.8; DB 4; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.3e-58;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG 60

Qy 61 CTGAAGAAGTGGATCGATTGTTGAGAAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTGAGAAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTAAGAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTAAGAAAATAGGAATAAA 180
```

```
; ORGANISM: Bacillus amyloliquefaciens
; US-09-672-459-3

Query Match          97.5%; Score 242.8; DB 4; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.3e-58;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG 60

Qy 61 CTGAAGAAGTGGATCGATTGTTGAGAAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTGAGAAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTAAGAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTAAGAAAATAGGAATAAA 180

Qy 181 GGGGGTGTGTTATTTACTGATATGTAAATATAATTTGTATAAGAAAATGAGAGGG 240
Db 181 GGGGGTGTGTTATTTACTGATATGTAAATATAATTTGTATAAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 15
US-09-636-252A-3
; Sequence 3, Application US/09636252A
; Patent No. 6440716
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: Alpha-Amylase Mutants
; FILE REFERENCE: 0776/1F216-US2
; CURRENT APPLICATION NUMBER: US/09/636,252A
; CURRENT FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 08/683,838
; PRIOR FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: B. amyloliquefaciens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (250)...(1791)
; NAME/KEY: mat_peptide
; LOCATION: (343)...(1791)
; NAME/KEY: sig_peptide
; LOCATION: (250)...(342)
; US-09-636-252A-3

Query Match          97.5%; Score 242.8; DB 4; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.3e-58;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG 60

Qy 61 CTGAAGAAGTGGATCGATTGTTGAGAAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTGAGAAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTAAGAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTAAGAAAATAGGAATAAA 180
```

QY 181 GGGGGTTCTTATTATTTTACTGATATGTAATAATATAATTTGTATATAAGAAAAATGAGAGG 240
Db 181 GGGGGTTCTTATTATTTTACTGATATGTAATAATATAATTTGTATATAAGAAAAATGAGAGG 240
QY 241 AGAGGA 246
Db 241 AGAGGA 246

Search completed: February 21, 2003, 08:06:42
Job time : 80 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 06:31:25 ; Search time 227 Seconds
(without alignments)
2470.254 Million cell updates/sec

Title: US-09-936-145-1

Perfect score: 249

Sequence: 1 gccccgacatacgaaga.....aaatgagaggagagatcc 249

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID52/qcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID52/qcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID52/qcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID52/qcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID52/qcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID52/qcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID52/qcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID52/qcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID52/qcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID52/qcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID52/qcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID52/qcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID52/qcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID52/qcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID52/qcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID52/qcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID52/qcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID52/qcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID52/qcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID52/qcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID52/qcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID52/qcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID52/qcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID52/qcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249	100.0	249	21	AAA93677
2	249	100.0	270	21	AAA93678
3	246	98.8	2604	20	AAE21079
4	246	98.8	2604	22	AAA37850
5	242.8	97.5	537	13	AAO22579
6	242.8	97.5	576	3	AAO20046
7	242.8	97.5	2084	16	AAQ88067
8	242.8	97.5	2084	16	AAO95032
9	242.8	97.5	2084	20	AA559681

10	242.8	97.5	2084	20	AA57596	Wild type Termamyl
11	242.8	97.5	2084	21	AAA48484	Bacillus amyloliqu
12	242.8	97.5	2084	24	ABL96211	Termamyl-like-alph
13	242.8	97.5	2084	24	ABL50568	B. amyloliquefacie
14	242.8	97.5	2084	24	AAI72215	Bacillus alpha amy
15	242.8	97.5	2084	24	AA520026	Bacillus DNA encod
16	230.8	92.7	2083	18	AAV02472	Bacillus amyloliqu
17	170.8	68.6	2166	24	AA344433	C-terminal Subtilli
18	170.8	68.6	2166	24	ABL40498	DNA construct A co
19	170.8	68.6	2267	24	AA344434	Subtilisin 309-CI-
20	170.8	68.6	2267	24	ABL40499	DNA construct B co
21	170.8	68.6	2588	24	ABL44436	Famyl AIG-barley C
22	170.8	68.6	2588	24	ABL40501	DNA construct D co
23	167.6	67.3	10216	17	AAT39279	Transposon deliver
24	162	65.1	185	20	AA23322	B. thuringiensis w
25	162	65.1	185	22	AAO9922	Bacillus amyloliqu
26	162	65.1	185	22	AA52645	Wild type amyQ pro
27	160.4	64.4	185	20	AA23325	Bacillus sp. amyQ
28	160.4	64.4	185	22	AAO9931	Bacillus amyloliqu
29	158.8	63.8	162	17	AAT85631	BAN promoter. Bac
30	157.2	63.1	185	20	AA23324	B. amyloliquefacie
31	157.2	63.1	185	22	AAO9924	Bacillus amyloliqu
32	157.2	63.1	185	22	AA52647	Consensus amyQ pro
33	155.6	62.5	185	20	AA23323	B. amyloliquefacie
34	155.6	62.5	185	22	AAO9923	Bacillus amyloliqu
35	155.6	62.5	185	22	AA52646	Consensus amyQ pro
36	107	43.0	6837	24	AA29902	Plasmid pMOL1642.
37	96.8	38.9	13222	22	AA25109	Nucleotide sequenc
38	46.4	18.6	201	6	AA50418	Sequence of the ar
39	40.4	16.2	48	22	AAO9917	Oligo #2 used in t
40	39.2	15.7	44	20	AA23340	Bacillus amyloliqu
41	39.2	15.7	44	22	AAO9916	Oligo #1 used in t
42	39.2	15.7	44	22	AA52640	amyQ promoter olig
43	39.2	15.7	44	22	AAI9490	Human neuroblastom
44	38.2	15.3	6887	24	ABN80014	Human chemically m
45	38.2	15.3	11155	24	ABL32604	Human immune syste

ALIGNMENTS

RESULT 1

AAA93677

ID AAA93677 standard; DNA; 249 BP.

AC AAA93677;

XX AAA93677;

DT 16-JAN-2001 (first entry)

XX Bacillus amyloliquefaciens alpha-amylase promoter.

DE Alpha-amylase promoter; restriction enzyme cleavage site;

XX expression cassette; expression vector; recombinant protein production;

KW ds.

OS Bacillus amyloliquefaciens.

XX WO200053778-A1.

PN 14-SEP-2000.

XX 08-MAR-2000; 2000WO-JP01415.

PF 08-MAR-1999; 99JP-0060904.

PR 06-OCT-1999; 99JP-0286034.

XX (SHOS) SHOWA SANGYO CO.

PI Inoue Y, Fushimi N, Mizubuchi H, Yamamoto Y, Ohshima Y;

PI Yasutake N, Miyoshi S;

DR WPI; 2000-594327/56.

XX

PT Modified Bacillus alpha-amylase promoter having additional restriction
 PT sites near the 3'-terminus for higher promotion of gene expression in
 PT Bacillus -
 XX
 XX
 PS Example 2; Page 39; 48pp; Japanese.
 XX
 CC The invention relates to a modified Bacillus amyloliquefaciens
 CC alpha-amylase promoter (AAA93678) which has a higher activity than the
 CC wild-type alpha-amylase promoter (AAA93677). The modified promoter
 CC contains several restriction enzyme cleavage sites near the 3' end.
 CC The invention also encompasses an expression cassette comprising the
 CC novel promoter, an expression vector comprising the expression cassette,
 CC host cells transformed with the expression vector, and the preparation
 CC of a gene product using cells transformed with the expression vector.
 CC The modified promoter of the invention and constructs and host cells
 CC comprising it may be used for the recombinant production of proteins
 CC for use as pharmaceuticals or foodstuffs. The present sequence
 CC represents the wild-type Bacillus amyloliquefaciens alpha-amylase
 CC promoter used in an exemplification of the invention.
 XX
 SQ Sequence 249 BP; 85 A; 33 C; 61 G; 70 T; 0 other;

Query Match 100.0%; Score 249; DB 21; Length 249;
 Best Local Similarity 100.0%; Pred. No. 8e-58;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTGG 60
 DB 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTGG 60
 QY 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAAGAGAGACCATAAAAATACCTTGTCTGTC 120
 DB 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAAGAGAGACCATAAAAATACCTTGTCTGTC 120
 QY 121 ATCAGACAGGGTATTTTATGCTGTCACACACTGTCGGCTGTGTAAGAAATAGAGATAAA 180
 DB 121 ATCAGACAGGGTATTTTATGCTGTCACACACTGTCGGCTGTGTAAGAAATAGAGATAAA 180
 QY 181 GGGGGTGTATTATTTTACTGATATGTAATAATAATTTGTAAGAAATAGAGAGG 240
 DB 181 GGGGGTGTATTATTTTACTGATATGTAATAATAATTTGTAAGAAATAGAGAGG 240
 QY 241 AGAGGATCC 249
 DB 241 AGAGGATCC 249

RESULT 2
 AAA93678
 ID AAA93678 standard; DNA; 270 BP.
 XX
 XX AAA93678;
 XX
 DT 16-JAN-2001 (first entry)
 XX
 DE Modified Bacillus amyloliquefaciens alpha-amylase promoter.
 XX
 KW Modified alpha-amylase promoter; restriction enzyme cleavage site;
 KW expression cassette; expression vector; recombinant protein production;
 KW mutant; ds.
 XX
 OS Bacillus amyloliquefaciens.
 OS Synthetic.
 XX
 PN W0200053778-A1.
 XX
 PD 14-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-JP01415.
 PF
 XX 08-MAR-1999; 99JP-0060904.
 PR
 XX 06-OCT-1999; 99JP-0286034.
 XX

PA (SHOS) SHOWA SANGYO CO.
 XX
 PI Inoue Y, Fushimi N, Mizubuchi H, Yamamoto Y, Onshima Y;
 PI Yasutake N, Miyoshi S;
 XX
 XX WPI; 2000-594327/56.
 DR
 XX Modified Bacillus alpha-amylase promoter having additional restriction
 PT sites near the 3'-terminus for higher promotion of gene expression in
 PT Bacillus -
 XX
 PS Claim 7; Page 39; 48pp; Japanese.
 XX
 CC This sequence represents a novel modified Bacillus amyloliquefaciens
 CC alpha-amylase promoter which has a higher activity than the
 CC wild-type alpha-amylase promoter (AAA93677). The modified promoter
 CC contains several restriction enzyme cleavage sites near the 3' end.
 CC The invention also encompasses an expression cassette comprising the
 CC novel promoter, an expression vector comprising the expression cassette,
 CC host cells transformed with the expression vector, and the preparation
 CC of a gene product using cells transformed with the expression vector.
 CC The modified promoter of the invention and constructs and host cells
 CC comprising it may be used for the recombinant production of proteins
 CC for use as pharmaceuticals or foodstuffs.
 XX
 SQ Sequence 270 BP; 89 A; 40 C; 67 G; 74 T; 0 other;

Query Match 100.0%; Score 249; DB 21; Length 270;
 Best Local Similarity 100.0%; Pred. No. 8.2e-58;
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTGG 60
 DB 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTGG 60
 QY 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAAGAGAGACCATAAAAATACCTTGTCTGTC 120
 DB 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAAGAGAGACCATAAAAATACCTTGTCTGTC 120
 QY 121 ATCAGACAGGGTATTTTATGCTGTCACACACTGTCGGCTGTGTAAGAAATAGAGATAAA 180
 DB 121 ATCAGACAGGGTATTTTATGCTGTCACACACTGTCGGCTGTGTAAGAAATAGAGATAAA 180
 QY 181 GGGGGTGTATTATTTTACTGATATGTAATAATAATTTGTAAGAAATAGAGAGG 240
 DB 181 GGGGGTGTATTATTTTACTGATATGTAATAATAATTTGTAAGAAATAGAGAGG 240
 QY 241 AGAGGATCC 249
 DB 241 AGAGGATCC 249

RESULT 3
 AAZ21079
 ID AAZ21079 standard; DNA; 2604 BP.
 XX
 XX AAZ21079;
 XX
 DT 18-NOV-1999 (first entry)
 XX
 DE Bacillus amyloliquefaciens Termamyl-like alpha-amylase encoding DNA.
 XX
 KW Alpha-amylase; Termamyl-like alpha-amylase; glucose syrup; starch; ss.
 XX
 OS Bacillus amyloliquefaciens.
 XX
 FH Key Location/Qualifiers
 FT -10_signal 707..712
 FT /*tag= a
 FT -35_signal 729..734
 FT /*tag= b
 FT RBS 759..762
 FT /*tag= c


```
FT CDS 770..2314
FT /*tag= d
FT /product= "Termamyl-like alpha-amylase"
FT sig_peptide 770..862
FT /*tag= e
FT mat_peptide 863..2311
FT /*tag= f
FT terminator 2321..2376
FT /*tag= g
FT
FT WO9946399-A1.
FT
FT 16-SEP-1999.
FT
FT 08-MAR-1999; 99WO-DK00114.
FT
FT 09-MAR-1998; 98DK-0000321.
FT
FT (NOVO ) NOVO-NORDISK AS.
FT
FT Norman BE, Hendriksen HV;
FT
FT WPI; 1999-551422/46.
FT P-PSDB; RAY29853.
FT
FT Preparation of a glucose syrup, using a Termamyl-like alpha-amylase -
FT
FT Disclosure; Page 29-32; 36pp; English.
FT
FT A method has been developed for the preparation of a glucose syrup using
FT a Termamyl-like alpha-amylase containing a substitution at Val(54).
FT The glucose syrup obtained by the process is useful as an ingredient in
FT food products. The Termamyl-like alpha-amylase facilitates the
FT preparation of glucose syrups suitable for the food industry, previously
FT only possible using acid hydrolysis. The present sequence encodes
FT Bacillus amyloliquefaciens Termamyl-like alpha-amylase.
FT
FT Sequence 2604 BP; 755 A; 516 C; 684 G; 649 T; 0 other;
FT
FT Query Match 98.8%; Score 246; DB 20; Length 2604;
FT Best Local Similarity 100.0%; Pred. No. 9.3e-57;
FT Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATGTAGCCCTTTTGATGACTGATGATTGG 60
Db 521 GCCCGGCACATACGAAAGACTGGCTGAAACATGTAGCCCTTTTGATGACTGATGATTGG 580
Qy 61 CTGAAGAAGTGCATGATGTTTGTAGAAAGAAAGAACCAATACCTTGTCTGTC 120
Db 581 CTGAAGAAGTGCATGATGTTTGTAGAAAGAAAGAACCAATACCTTGTCTGTC 640
Qy 121 ATCAGACAGGCTATTTTATGCTGCCAGACTGTCGCTGTGTAATAAATAGGAATAA 180
Db 641 ATCAGACAGGCTATTTTATGCTGCCAGACTGTCGCTGTGTAATAAATAGGAATAA 700
Qy 181 GGGGGTGTGTTATTTACTGATATGATAAATATAATTTGTATAGAATAATGAGAGGG 240
Db 701 GGGGGTGTGTTATTTACTGATATGATAAATATAATTTGTATAGAATAATGAGAGGG 760
Qy 241 AGAGGA 246
Db 761 AGAGGA 766
FT
FT RESULT 4
FT AAA37850
FT ID AAA37850 standard; DNA; 2604 BP.
FT
FT AC AAA37850;
FT
FT 12-FEB-2001 (first entry)
FT
FT B. amyloliquefaciens termamyl-like alpha amylase coding sequence.
```

```
XX Termamyl-like alpha-amylase; variant; starch liquefaction; fuel;
KW detergent composition; laundry cleaning composition; ethanol production;
KW dish washing cleaning composition; hard surface cleaning composition;
KW industrial ethanol production; textile desizing; ds.
OS Bacillus amyloliquefaciens.
XX
FH Key Location/Qualifiers
FT -10_signal 707..712
FT /*tag= a
FT -35_signal 729..734
FT /*tag= b
FT RBS 759..762
FT /*tag= c
FT CDS 770..2314
FT /*tag= d
FT /product= alpha-amylase
FT /note= "the signal peptide is not shown in the encoded
FT protein given in the specification"
FT sig_peptide 770..862
FT /*tag= e
FT /note= "the signal peptide is not shown in the
FT specification"
FT
FT mat_peptide 863..2311
FT /*tag= f
FT terminator 2321..2376
FT /*tag= g
FT
FT WO200060059-A2.
XX
PN 12-OCT-2000.
XX
XX 28-MAR-2000; 2000WO-DK00148.
XX
XX 30-MAR-1999; 99DK-0000437.
XX
XX (NOVO ) NOVO NORDISK AS.
XX
XX Andersen C, Jorgensen CT, Bisgard-Frantzen H, Svendsen A;
XX Kjaerulff S;
XX WPI; 2001-015656/02.
XX P-PSDB; RAY97546.
XX
XX New variants of parent Termamyl-like alpha-amylase, useful in starch
XX liquefaction, in detergent compositions and in ethanol production,
XX exhibit altered cleavage pattern relative to the parent -
XX
XX Disclosure; Page 64-67; 78pp; English.
XX
XX This sequence encodes a termamyl-like alpha amylase.
XX The invention relates to a variant (I) of parent Termamyl-like
XX alpha-amylase comprising alteration at one or more of the positions
XX W13, G48, T49, S50, Q51, A52, D53, V54, G57, G107, G108, A111, S168 and
XX M197. The alterations in (I) are independently an insertion of an amino
XX acid downstream of the amino acid which occupies the position or deletion
XX or substitution of the amino acid which occupies the position with a
XX different amino acid. The variant has alpha-amylase activity. (I) or
XX compositions containing it are useful in starch liquefaction, in
XX detergent composition such as laundry, dish washing and hard surface
XX cleaning compositions, ethanol production such as fuel, drinking and
XX industrial ethanol production, desizing of textiles, fabrics or garments.
XX (I) exhibits a reduced capability of cleaving a substrate close to the
XX branching point, and further exhibits improved substrate specificity
XX and/or improved specific activity relative to the parent alpha-amylase.
XX
XX Sequence 2604 BP; 755 A; 516 C; 684 G; 649 T; 0 other;
XX
XX Query Match 98.8%; Score 246; DB 22; Length 2604;
XX Best Local Similarity 100.0%; Pred. No. 9.3e-57;
XX Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTGG 60
Db 521 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTGG 580
QY 61 CTGAAGAAGTGCATGATTGTTTGAGAAAAGAGAAGACCAATAAAATACCTTCTCTGTC 120
Db 581 CTGAAGAAGTGCATGATTGTTTGAGAAAAGAGAAGACCAATAAAATACCTTCTCTGTC 640
QY 121 ATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAATAAGGAATAAAA 180
Db 641 ATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAATAAGGAATAAAA 700
QY 181 GGGGGTGTATTATTTACTGATATGTAATAATATATTTGTAAGAAAATGAGAGGG 240
Db 701 GGGGGTGTATTATTTACTGATATGTAATAATATATTTGTAAGAAAATGAGAGGG 760
QY 241 AGAGGA 246
Db 761 AGAGGA 766

RESULT 5
AAQ22579
ID AAQ22579 standard; DNA; 537 BP.
XX
AC AAQ22579;
XX
DT 23-JUL-1992 (first entry)
XX
DE Sequence of the NH2 region of the B. subtilis alpha-amylase gene
DE contg. new ClaI sites.
XX
KW Heterologous regulatory promoter module; hyperexpression; secretion;
KW Bacillus expression system; RNA polymerase initiation site; ds.
XX
OS Bacillus amyloliquefaciens.
XX
FH Key Location/Qualifiers
FT -35_signal 193..201 /*tag= a
FT -10_signal 209..214 /*tag= b
FT sig_peptide 250..242 /*tag= c
FT misc_feature 119 /*tag= d
FT /*note= "ClaI site introd. by primer 302"
FT misc_feature 144 /*tag= e
FT /*note= "ClaI site introd. by primer 303"
FT misc_feature 172 /*tag= f
FT /*note= "ClaI site introd. by primer 304"
XX
PN WO9203561-A.
XX
XX 05-MAR-1992.
XX
PF 12-AUG-1991; 91WO-F100244.
XX
PR 13-AUG-1990; 90US-0565847.
XX
PA (ALKO-) ALKO LTD.
XX
PI Palva IA, Palva AM;
XX
DR WPI; 1992-096904/12.
XX
DR P-PSDB; AAR21689.
XX
PT Hybrid alpha-amylase promoter for recombinant gene expression -
PT in prokaryote(s), esp. Bacillus subtilis, contg. target module of
PT enhancer gene for ribonucleic acid
XX
```

```
PS Example; Fig 3; 51pp; English.
XX
CC The inventors claim a new hybrid promoter for the expression of
CC recombinant genes in a prokaryotic host. The promoter comprises a
CC target module of an enhancer gene operably linked to an initiation
CC module for RNA polymerase. The target module is found in the 5',
CC regulatory region of the apr gene, esp. from B. subtilis. It is a
CC target module of a protein prod. of an enhancer gene esp. the B.
CC subtilis sacd enhancer gene, the sacd enhancer gene of B. subtilis,
CC B. amyloliquefaciens or B. licheniformis, or the prtr enhancer gene
CC of B. netto or B. subtilis. The initiation module is that of a gene
CC encoding a prokaryotic exoenzyme, esp. B. amyloliquefaciens alpha-
CC amylase (AA).
XX
SQ Sequence 537 BP; 171 A; 90 C; 129 G; 147 T; 0 other;
Query Match 97.5%; Score 242.8; DB 13; Length 537;
Best Local Similarity 99.2%; Pred. No. 4.6e-56;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTGG 60
QY 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAAGACCAATAAAATACCTTCTCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAAGACCAATAAAATACCTTCTCTGTC 120
QY 121 ATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAATAAGGAATAAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGTCAGACTGTCGCTGTGTAAATAAGGAATAAAA 180
QY 181 GGGGGTGTATTATTTACTGATATGTAATAATATATTTGTAAGAAAATGAGAGGG 240
Db 181 GGGGGTGTATTATTTACTGATATGTAATAATATATTTGTAAGAAAATGAGAGGG 240
QY 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 6
AAN20046
ID AAN20046 standard; DNA; 576 BP.
XX
AC AAN20046;
XX
DT 09-OCT-1992 (first entry)
XX
DE Bacillus amyloliquefaciens alpha-amylase gene.
XX
KW Recombinant protein production; secretory peptide;
KW signal sequence; protein secretion; ds.
XX
OS Bacillus amyloliquefaciens.
XX
PN BE891659-A.
XX
PD 16-APR-1982.
XX
PF 30-DEC-1981; 81BE-0891659.
XX
PR 31-DEC-1980; 80FI-0004081.
PR 12-MAR-1982; 82FI-0000860.
XX
PA (PALV/) PALVA I.
PA (GENE-) GENESIT Oy.
XX
DR WPI; 1982-37323E/19.
XX
PT Protein prodn. using bacillus strain bacteria - by splitting the
PT alpha-amylase bacillus gene, combining the DNA coding for the
PT protein and joining to a plasmid in the bacillus
XX
```

```

XX PS Claim 9; Page 25; 38pp; French.
XX CC Genomic DNA was isolated from B.amyloliquefaciens and digested with
XX CC MboI. DNA fragments of 1.5-5.0 x 10power6 daltons were ligated to
XX CC plasmid pUB110 (isolated from B.subtilis SB202) which had been
XX CC linearised with BamHI. Recombinant plasmids were used to transform
XX CC B.subtilis 1A197 (contg. mutant alpha-amylase gene amy-).
XX CC Transformants containing the alpha-amylase coding sequence were
XX CC identified and the gene was isolated and sequenced. The secretory
XX CC signal sequence from the B.amyloliquefaciens alpha-amylase gene
XX CC was used in the construction of vectors which are useful for
XX CC efficient production of proteins in Bacillus subtilis hosts;
XX CC the alpha-amylase signal peptide from B.amyloliquefaciens is 10 x
XX CC more efficient than that from B.subtilis. the preferred sequence
XX CC for use in such vectors covers nucleotides 306 to 349 of AAN20046.
XX CC See also AAN20145-N20154.
XX SQ Sequence 576 BP; 190 A; 98 C; 139 G; 149 T; 0 other;

  Query Match          97.5%; Score 242.8; DB 3; Length 576;
  Best Local Similarity 99.2%; Pred. No. 4.6e-56;
  Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTTGAGCTTTGATGACTGATGATTGG 60
DB 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3 GCCCGGCACATACGAAAGACTGGCTGAAACATTTGAGCTTTGATGACTGATGATTGG 62
DB 3 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 CTGAAGAAAGTGGATCGATTGTTTGAGAAAGAGAACACCAATAAAATACCTTGTCTGC 120
DB 61 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 63 CTGAAGAAAGTGGATCGATTGTTTGAGAAAGAGAACACCAATAAAATACCTTGTCTGC 122
DB 63 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGTTGTAATAAATAGGAATAAA 180
DB 121 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 123 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGTTGTAATAAATAGGAATAAA 182
DB 123 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GGGGGTGTGTTATTTTACTGATATCTGTAATAATATATTTGTAAGAAATGAGAGGG 240
DB 181 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 183 GGGGGTGTGTTATTTTACTGATATCTGTAATAATATATTTGTAAGAAATGAGAGGG 242
DB 183 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 AGAGGA 246
DB 241 |||||
QY 243 AGAGGA 248
DB 243 |||||

RESULT 7
ID AAQ88067 standard; DNA; 2084 BP.
XX AC AAQ88067;
XX DT 01-DEC-1995 (first entry)
XX DE Bacillus amyloliquefaciens alpha amylase coding sequence.
XX KW Alpha amylase; variant; enzyme; detergent; additive; dishwashing;
XX KW washing; Bacillus licheniformis; Bacillus amyloliquefaciens;
XX KW Bacillus stearothermophilus; dyeing; bleaching; scouring; textile;
XX KW thermostable; ss.
XX OS Bacillus amyloliquefaciens.
XX XX
XX FH Key Location/Qualifiers
XX FT CDS 250..1794
XX FT /*tag= a
XX FT /product= Alpha amylase.
XX FT sig_peptide 250..342
XX FT /*tag= b
XX FT mat_peptide 343..1791
XX FT /*tag= c
XX PN W09510603-A.

```

```

PD 20-APR-1995.
XX PF 05-OCT-1994; 94WO-DK00370.
XX PR 08-OCT-1993; 93DK-0001133.
XX PR 02-FEB-1994; 94DK-0000140.
XX PA (NOVO ) NOVO-NORDISK AS.
XX PI Svendsen A, Thellersen M;
XX PI Van der zee P, Bisgardfrantzen H, Borchert T;
XX DR WPT: 1995-161790/21.
XX DR P-FSDB; AAR72448.
XX PT New Bacillus derived alpha-amylase variants - having amino acid
XX PT modifications to improve washing and/or dishwashing performance
XX PS Disclosure; Page 72-73; 105pp; English.
XX CC Variant alpha amylase enzymes which have improved washing and/or
XX CC as detergent additives. The enzymes have one or more amino acid
XX CC residues added, deleted or substituted. The variants can also be
XX CC used for textile desizing prior to scouring, bleaching and dyeing.
XX CC The variants have improved thermostability, acid/alkaline stability;
XX CC low temperature optimum; pH optimum; higher hydrolysis velocity and
XX CC improved tolerance to other composition constituents, e.g. oxidation
XX CC agents.
XX SQ Sequence 2084 BP; 610 A; 401 C; 544 G; 529 T; 0 other;

  Query Match          97.5%; Score 242.8; DB 16; Length 2084;
  Best Local Similarity 99.2%; Pred. No. 6.4e-56;
  Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTTGAGCTTTGATGACTGATGATTGG 60
DB 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1 GCGCCGCACATACGAAAGACTGGCTGAAACATTTGAGCTTTGATGACTGATGATTGG 60
DB 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 CTGAAGAAAGTGGATCGATTGTTTGAGAAAGAGAACACCAATAAAATACCTTGTCTGC 120
DB 61 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGTTGTAATAAATAGGAATAAA 180
DB 121 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 GGGGGTGTGTTATTTTACTGATATCTGTAATAATATATTTGTAAGAAATGAGAGGG 240
DB 181 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 241 AGAGGA 246
DB 241 |||||
QY 243 AGAGGA 248
DB 243 |||||

RESULT 8
ID AAQ95032 standard; DNA; 2084 BP.
XX AC AAQ95032;
XX DT 17-JAN-1996 (first entry)
XX DE Bacillus amyloliquefaciens alpha amylase gene.
XX KW Alpha amylase; oxidation; desizing; bleaching; scouring; fabric;
XX KW starch; thermostable; methionine; Bacillus licheniformis;
XX KW Bacillus amyloliquefaciens; Bacillus stearothermophilus; ss.
XX OS Bacillus amyloliquefaciens.
XX XX
XX FH Key Location/Qualifiers

```

FT	5' UTR	1..249	
FT		/*tag= a	
FT	CDS	250..1794	
FT		/*tag= b	
FT		/product= Alpha amylase.	
FT	sig_peptide	250..342	
FT		/*tag= c	
FT	mat_peptide	343..1791	
FT		/*tag= d	
FT	3' UTR	1795..2084	
FT		/*tag= e	
XX			
PN	W09521247-Al.		
XX			
XX	10-AUG-1995.		
XX			
XX	05-OCT-1994;	94WO-DK00371.	
XX			
XX	02-FEB-1994;	94DK-0000141.	
PR			
XX	(NOVO) NOVO-NORDISK AS.		
PA			
PA	Marcher D, Nilsson TE, Pedersen HH, Toft AH;		
PI			
XX			
XX	WPI; 1995-283767/37.		
DR	P-PSDB; AAR78268.		
XX			
XX	use of an oxidation stable alpha-amylase - for simultaneous desizing		
PT	and bleaching or scouring of fabrics contg. starch or starch derivs.		
PT			
XX	Disclosure: Page 22-24; 37pp; English.		
XX			
CC	Oxidation stable alpha amylases can be used for the simultaneous		
CC	desizing and bleaching or scouring of a fabric comprising starch or		
CC	starch derivatives. They exhibit a better heat stability,		
CC	especially in the presence of oxidising agents. They are obtained		
CC	from a parent alpha amylase by replacing one or more methionine		
CC	residues with any amino acid different from Cys or Met, preferably		
CC	Leu, Thr, Ala, Gly, Ser, Ile or Asp. The parent alpha amylase is		
CC	derived from a Bacillus species. This sequence encodes the wild		
CC	type (unmodified) alpha amylase.		
XX			
SQ	Sequence 2084 BP; 610 A; 401 C; 544 G; 529 T; 0 other;		
	Query Match	97.5%; Score 242.8; DB 16;	
	Best Local Similarity	99.28; Pred. No. 6.4e-56;	
	Matches 244; Conservative	0; Mismatches 2; Indels 0; Gaps	
Qy	1	GCCTCCGCACATACGAAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG	60
Db	1	GCCTCCGCACATACGAAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG	60
Qy	61	CTGAGAGAGTGGATCGATTGTTTGAGAAAAAGAGAAGACCATATAAAATACCTTGTCTGTC	120
Db	61	CTGAGAGAGTGGATCGATTGTTTGAGAAAAAGAGAAGACCATATAAAATACCTTGTCTGTC	120
Qy	121	ATCAGACAGGGTATTTTTTATGTCTGCCAGACTGTCCGCTGTGTAAAAAATAGGAATAAA	180
Db	121	ATCAGACAGGGTATTTTTTATGTCTGCCAGACTGTCCGCTGTGTAAAAAATAGGAATAAA	180
Qy	181	GGGGGGTTCTTATATTTTACTGATATGTAAAAATATAAATTTCTATAAGAAAATGAGAGGG	240
Db	181	GGGGGGTTCTTATATTTTACTGATATGTAAAAATATAAATTTCTATAAGAAAATGAGAGGG	240
Qy	241	AGAGGA 246	
Db	241	AGAGGA 246	
RESULT 9			
AAK59681			
ID	AAK59681 standard; DNA; 2084 BP.		
XX			


```
Db 121 ATCAGACAGGGTATTTTATGCTGCTCAGACTGCTCGCTGTGTAAATAAAGGAATAAA 180
Qy 181 GGGGGTGTATTATTTACTGATATGTAATAATATTTTGTATATAAGAAAATCAGAGG 240
Db 181 GGGGGTGTATTATTTACTGATATGTAATAATATTTTGTATATAAGAAAATCAGAGG 240
Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 12
ID ABL96211 standard; DNA; 2084 BP.
AC ABL96211;
XX
DT 19-AUG-2002 (first entry)
XX
DE Termamyl-like-alpha-amylases encoding sequence #5.
XX
DE Termamyl; alpha amylase; starch liquefaction; ethanol production;
KW textile desizing; detergent; enzyme; gene; ds.
XX
OS Bacillus amyloliquefaciens.
XX
FH Key Location/Qualifiers
FT CDS 343..1794
FT /*tag= a
FT /product= "termamyl-like alpha amylase"
FT /partial
FT /note= "no start codon"
XX
XX WO200210355-A2.
XX
XX 07-FEB-2002.
PD
XX
PF 12-JUL-2001; 2001WO-DK00488.
XX
XX 01-AUG-2000; 2000DK-0001160.
PR
XX 12-SEP-2000; 2000DK-0001354.
PR
XX 10-NOV-2000; 2000DK-0001887.
PR
XX 26-APR-2001; 2001DK-0000655.
XX
XX (NOVO ) NOVOZYMES AS.
XX
XX Thisted T, Kjaerulff S, Andersen C, Fuglsang CC;
PI WPI; 2002-280633/32.
XX
XX P-PSDB; ABB76590.
DR
XX
PT Variant of parent Termamyl-like alpha amylase, useful in detergent
PT compositions, for starch liquefaction, ethanol production, washing
PT and/or dish washing, and textile desizing -
XX
XX Disclosure; Page 69-71; 90pp; English.
PS
XX
CC This invention relates to variants of a parent Termamyl-like
CC alpha-amylases. These are used for starch liquefaction, ethanol
CC production, detergent, and textile desizing. The amylases have altered
CC stability, particularly at high temperatures from 70-120plusoc and
CC low pH in the range from pH 4.0-6.0. The present sequence is a
CC termamyl-like-alpha-amylase encoding sequence.
XX
XX Sequence 2084 BP; 610 A; 401 C; 544 G; 529 T; 0 other;
SQ
Query Match 97.5%; Score 242.8; DB 24; Length 2084;
Best Local Similarity 99.2%; Pred. No. 6.4e-56;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GCCCGGCACATACGAAGACTGCTGCAACATTTGAGCCTTTGATGACTGATGATTGG 60
|||||
```

```
Db 1 GCCCGGCACATACGAAGACTGCTGCAACATTTGAGCCTTTGATGACTGATGATTGG 60
Qy 61 CTGAAGAAGTGGATCGATGTTTGTAGAAAAAGAGAGACCATAAAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGGATCGATGTTTGTAGAAAAAGAGAGACCATAAAAATACCTTGTCTGTC 120
Qy 121 ATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGGCTGTGTAAAAAATAGGAATAAA 180
Qy 181 GGGGGTGTATTATTTACTGATATGTAATAATATTTTGTATATAAGAAAATGAGAGG 240
Db 181 GGGGGTGTATTATTTACTGATATGTAATAATATTTTGTATATAAGAAAATGAGAGG 240
Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 13
ABL50568
ID ABL50568 standard; DNA; 2084 BP.
XX
AC ABL50568;
XX
DT 19-JUN-2002 (first entry)
XX
DE B. amyloliquefaciens termamyl-like alpha-amylase encoding DNA SEQ ID:9.
XX
DE Bacillus; termamyl-like alpha-amylase; alpha-amylase; EC 3.2.1.1;
KW variant; mutant; enzyme; protein co-ordinate data; cleaning; detergent;
KW washing; sweetener; ethanol; starch; gene; ds.
XX
XX Bacillus amyloliquefaciens.
XX
XX Key Location/Qualifiers
FT CDS 343..1794
FT /*tag= a
FT /EC_number= "3.2.1.1"
FT /product= "termamyl-like alpha-amylase"
XX
XX WO200166712-A2.
XX
XX 13-SEP-2001.
PD
XX
XX 07-MAR-2001; 2001WO-DK00144.
XX
XX 08-MAR-2000; 2000DK-0000376.
PR
XX 15-MAR-2000; 2000US-189857P.
PR
XX 23-FEB-2001; 2001DK-0000303.
PR
XX 26-FEB-2001; 2001US-271382P.
XX
XX (NOVO ) NOVOZYMES AS.
XX
XX Andersen C, Borchert TV, Nielsen BR;
PI WPI; 2002-239612/29.
XX
XX P-PSDB; ABB06937.
DR
XX
XX Novel variant of parent termamyl-like alpha-amylase useful as a
PT component in washing and dishwashing compositions, for textile
PT desizing, for starch liquefaction, and for producing sweeteners and
PT ethanol from starch -
XX
XX Disclosure; Page 143-145; 153pp; English.
PS
XX
CC The present invention describes a variant of a parent termamyl-like
CC alpha-amylase (EC 3.2.1.1) (I) comprising an alteration at one or more
CC positions of a group of 31 possible amino acid positions. The alteration
CC in (I) may be at Arg28, Arg118, Asn174, Arg181, Gly182, Asp183, Gly184,
CC Gly186, Trp189, Asn195, Met202, Tyr298, Asn299, Lys302, Asn306,
CC Arg310, Asn314, Arg320, His324, Glu345, Tyr396, Arg400, Trp439, Arg444,
CC Asn445, Lys446, Gln449, Arg458, Asn471, or Asn484. (I) can be used for
```

CC washing and/or dishwashing, textile desizing, and starch liquefaction.
 CC (I) is useful as a component in hard surface cleaning detergent
 CC composition, and for producing sweeteners and ethanol from starch.
 CC (I) has altered solubility, preferably increased solubility, in
 CC particular under washing, dish washing or hard surface cleaning
 CC conditions. The present sequence encodes a *Bacillus amyloliquefaciens*
 CC termamyl-like alpha-amylase which is used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 2084 BP; 610 A; 401 C; 544 G; 529 T; 0 other;
 Query Match 97.5%; Score 242.8; DB 24; Length 2084;
 Best Local Similarity 99.2%; Pred. No. 6.4e-56;
 Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTTGG 60
 Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTTGG 60
 Qy 61 CTGAAGAAGTGGATCGATTGTTGAGAAAGAAGAACACCATAAAATACCTTGTCTGTC 120
 Db 61 CTGAAGAAGTGGATCGATTGTTGAGAAAGAAGAACACCATAAAATACCTTGTCTGTC 120
 Qy 121 ATCAGACAGGCTATTTTATGCTGCTCCAGACTGTCGCCCTGTGTAATAAATAGGAATAA 180
 Db 121 ATCAGACAGGCTATTTTATGCTGCTCCAGACTGTCGCCCTGTGTAATAAATAGGAATAA 180
 Qy 181 GGGGGTGTATTATTTACTGATATGTAATAATATTTGTTATAGAAATAGAGAGG 240
 Db 181 GGGGGTGTATTATTTACTGATATGTAATAATATTTGTTATAGAAATAGAGAGG 240
 Qy 241 AGAGGA 246
 Db 241 AGAGGA 246
 RESULT 14
 AAI72215
 ID AAI72215 standard; cDNA; 2084 BP.
 XX
 AC AAI72215;
 XX
 DT 02-APR-2002 (first entry)
 XX
 DE *Bacillus alpha amylase* BAN cDNA.
 XX
 KW Alpha amylase; *Bacillus*; Termamyl-like; maltodextrin; glucose syrup;
 KW starch; food; feed; pharmaceutical; confectionery; candy;
 KW isotonic drink; bakery; cereal bar; ice cream; coffee whitener;
 KW salad dressing; cured meat; fermented meat; spice; ss.
 XX
 OS *Bacillus amyloliquefaciens*.
 XX
 FH Key Location/Qualifiers
 FT CDS 343..1794
 FT /*tag= a
 FT /product= "Alpha-amylase"
 XX
 PN WO200196537-A2.
 XX
 PD 20-DEC-2001.
 XX
 XX 13-JUN-2001; 2001WO-DK00404.
 XX
 PR 14-JUN-2000; 2000DK-0000917.
 PR
 PR 20-JUN-2000; 2000US-212852P.
 XX
 XX (NOVO) NOVOZYMES AS.
 XX
 XX Nielsen BR, Weibye M;
 PI
 XX WPI; 2002-098064/13.
 XX
 DR P-PSDB; AAB47854.

XX New modified alpha-amylase derived from the genus *Bacillus* and/or is a
 PT Termamyl-like alpha-amylase, which has been pre-oxidized for producing
 PT maltodextrin or glucose syrup
 XX
 PS Claim 7; Page 37-39; 47pp; English.
 XX
 CC The sequences given in AAI72211-16 encode modified alpha-amylases
 CC derived from the genus *Bacillus*. These alpha amylases are Termamyl-
 CC like alpha-amylase and they have been pre-oxidized. The alpha amylase
 CC is useful for producing a maltodextrin or glucose syrup, by treating
 CC starch with a pre-oxidized alpha-amylase until a product with a
 CC DE between 5-45 has been provided and/or until a product with a
 CC molecular weight of between 5-30 kda has been provided. The product
 CC comprises a maltodextrin with a DE of 18.5 and/or a maltodextrin with a
 CC molecular weight of 14-16 kda. The alpha amylase is useful for producing
 CC a maltodextrin or glucose syrup, where the glucose syrup is useful as an
 CC ingredient in food, feed or pharmaceuticals. Glucose syrup is useful
 CC in confectionery such as candies, beverages such as isotonic drinks,
 CC bakery such as cereal bars, dairy and ice cream such as coffee
 CC whiteners, conventional foods such as salad dressings, and food
 CC ingredients and preparations such as cured meat, fermented meat, spices
 CC and seasoning encapsulated flavours.
 XX
 SQ Sequence 2084 BP; 610 A; 401 C; 544 G; 529 T; 0 other;
 Query Match 97.5%; Score 242.8; DB 24; Length 2084;
 Best Local Similarity 99.2%; Pred. No. 6.4e-56;
 Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTTGG 60
 Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTTGG 60
 Qy 61 CTGAAGAAGTGGATCGATTGTTGAGAAAGAAGAACACCATAAAATACCTTGTCTGTC 120
 Db 61 CTGAAGAAGTGGATCGATTGTTGAGAAAGAAGAACACCATAAAATACCTTGTCTGTC 120
 Qy 121 ATCAGACAGGCTATTTTATGCTGCTCCAGACTGTCGCCCTGTGTAATAAATAGGAATAA 180
 Db 121 ATCAGACAGGCTATTTTATGCTGCTCCAGACTGTCGCCCTGTGTAATAAATAGGAATAA 180
 Qy 181 GGGGGTGTATTATTTACTGATATGTAATAATATTTGTTATAGAAATAGAGAGG 240
 Db 181 GGGGGTGTATTATTTACTGATATGTAATAATATTTGTTATAGAAATAGAGAGG 240
 Qy 241 AGAGGA 246
 Db 241 AGAGGA 246
 RESULT 15
 AAS20026
 ID AAS20026 standard; DNA; 2084 BP.
 XX
 AC AAS20026;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE *Bacillus* DNA encoding TERMAMYL-like alpha-amylase BAN.
 XX
 KW TERMAMYL; alpha-amylase; ds; detergent; dishwashing; textile desizing;
 KW starch liquefaction; ethanol production; hard surface cleaner;
 KW sweetener; amylopectin; limit dextrin; NOVAMYL; BAN.
 XX
 OS *Bacillus amyloliquefaciens*.
 XX
 FH Key Location/Qualifiers
 FT CDS 343..1794
 FT /*tag= a
 FT /product= "Alpha-amylase BAN"
 FT /partial
 FT /note= "NO start codon"

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 21, 2003, 06:32:14 ; Search time 1154 seconds
(without alignments)
6279.548 Million cell updates/sec

Title: US-09-936-145-1
Perfect score: 249
Sequence: 1 gccccgcacatacgaaga.....aaatgagaggagagatccc 249

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	246	98.8	2604	6	ARI68304	ARI68304 Sequence
2	246	98.8	2604	6	AX036892	AX036892 Sequence
3	242.8	97.5	537	1	BAAMYL	V00092 Bacillus am
4	242.8	97.5	576	6	A00155	A00155 Nucleotide
5	242.8	97.5	576	6	A00156	A00156 Nucleotide
6	242.8	97.5	576	6	A00604	A00604 B.amyloliqu
7	242.8	97.5	576	6	A00605	A00605 B.amyloliqu
8	242.8	97.5	2084	1	BACAM	J01542 Bacillus am
9	242.8	97.5	2084	6	AR008285	AR008285 Sequence
10	242.8	97.5	2084	6	AR037275	AR037275 Sequence
11	242.8	97.5	2084	6	AR052145	AR052145 Sequence
12	242.8	97.5	2084	6	AR087551	AR087551 Sequence
13	242.8	97.5	2084	6	ARI29916	ARI29916 Sequence
14	242.8	97.5	2084	6	ARI37905	ARI37905 Sequence
15	242.8	97.5	2084	6	ARI43262	ARI43262 Sequence
16	242.8	97.5	2084	6	AX244195	AX244195 Sequence
17	242.8	97.5	2084	6	AX305020	AX305020 Sequence
18	242.8	97.5	2084	6	AX339268	AX339268 Sequence
19	242.8	97.5	2084	6	AX370717	AX370717 Sequence
20	241.2	96.9	576	6	A00157	A00157 Nucleotide
21	241.2	96.9	576	6	A00158	A00158 Nucleotide
22	162	65.1	185	6	ARI60501	ARI60501 Sequence
23	162	65.1	185	6	AX088919	AX088919 Sequence
24	158.8	63.8	162	6	I40596	I40596 Sequence 24
25	158.8	63.8	162	6	I40867	I40867 Sequence 24
26	158.8	63.8	162	6	I40895	I40895 Sequence 24
27	158.8	63.8	162	6	I56837	I56837 Sequence 24
28	157.2	63.1	185	6	ARI60503	ARI60503 Sequence
29	157.2	63.1	185	6	AX088921	AX088921 Sequence
30	155.6	62.5	185	6	ARI60502	ARI60502 Sequence
31	155.6	62.5	185	6	AX088920	AX088920 Sequence
32	149	59.8	465	6	A00607	A00607 B.amyloliqu
33	149	59.8	465	6	A00608	A00608 B.amyloliqu
34	145.8	58.6	465	6	E00040	E00040 DNA coding
35	96.8	38.9	13222	6	AX167074	AX167074 Sequence
36	52	20.9	232	12	SYNPTUB261	M36664 Synthetic p
37	52	20.9	232	12	SYNPTUB263	M36665 Synthetic p
38	51.2	20.6	185	12	SYNBACAMY	M34931 Synthetic B
39	46.4	18.6	201	6	A01100	A01100 Nucleotide
40	46.4	18.6	201	6	A04541	A04541 Artificial
41	46.4	18.6	282	12	SYNAAAGFP	M35837 C.tetragono
42	44	17.7	13851	8	MICEGEN3	Y13644 Chlorogonlu
43	41	16.5	157069	9	AC023934	AC023934 Homo sapi
44	41	16.5	187599	9	AC021010	AC021010 Homo sapi
45	40.4	16.2	181062	2	AC062006	AC062006 Homo sapi

ALIGNMENTS

RESULT 1	ARI68304	ARI68304	Sequence 3 from patent US 6287826.	2604 bp	DNA	linear	PAT 17-DEC-2001
LOCUS	ARI68304	Sequence 3 from patent US 6287826.					
DEFINITION	ARI68304	Sequence 3 from patent US 6287826.					
ACCESSION	ARI68304	Sequence 3 from patent US 6287826.					
VERSION	ARI68304.1	GI:17904148					
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unknown.						
REFERENCE	1 (bases 1 to 2604)						
AUTHORS	Norman,B.Edmund. and Hendriksen,H.Vang.						
TITLE	Enzymatic preparation of glucose syrup from starch						
JOURNAL	Patent: US 6287826-A 3 11-SEP-2001;						
FEATURES	Location/Qualifiers						

BASE COUNT	termin
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Db 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAGAAAGAACCATATAAAATACCTTGCTCTGC 120

Qy 121 ATCAGACAGGGTATTTTATCTGCTCCAGACTGTCGCGTGTAAATAATAGGAATAAA 180
|||||
Db 121 ATCAGACAGGGTATTTTATCTGCTCCAGACTGTCGCGTGTAAATAATAGGAATAAA 180
|||||

Qy 181 GGGGGGTTGTTATTATTTTACTGATGTGTAATAATAATTTGTATAGAATAATGAGAGG 240
|||||
Db 181 GGGGGGTTGTTATTATTTTACTGATGTGTAATAATAATTTGTATAGAATAATGAGAGG 240
|||||

Qy 241 AGAGGA 246
|||||

Db 241 AGAGGA 246
|||||

RESULT 4
A00155

LOCUS A00155 576 bp DNA linear PAT 28-JAN-1993
DEFINITION Nucleotide sequence 1 from patent number GB2133408.
ACCESSION A00155
VERSION A00155.1 GI:14423
KEYWORDS
SOURCE Bacillus sp.
ORGANISM Bacillus sp.
REFERENCE 1 (bases 1 to 576)
AUTHORS Palva,I.
JOURNAL Patent: GB 2133408-A 1 25-JUL-1984;
Ikka Palva

FEATURES
source Location/Qualifiers
1..576
/db_xref="taxon:1409"

BASE COUNT 190 a 98 c 139 g 149 t

ORIGIN
Query Match 97.5%; Score 242.8; DB 6; Length 576;
Best Local Similarity 99.2%; Pred. No. 4.6e-48;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCGCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCCTTTGATGACTGATGATTGG 60
|||||
Db 574 GCGCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCCTTTGATGACTGATGATTGG 515
|||||

Qy 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAGAAAGAACCATATAAAATACCTTGCTCTGC 120
|||||
Db 514 CTGAAGAAGTGGATCGATTGTTTGAGAAAGAAAGAACCATATAAAATACCTTGCTCTGC 455
|||||

Qy 121 ATCAGACAGGGTATTTTATCTGCTCCAGACTGTCGCGTGTAAATAATAGGAATAAA 180
|||||
Db 454 ATCAGACAGGGTATTTTATCTGCTCCAGACTGTCGCGTGTAAATAATAGGAATAAA 395
|||||

Qy 181 GGGGGTGTATTATTATTTTACTGATGTGTAATAATAATTTGTATAGAATAATGAGAGG 240
|||||
Db 394 GGGGGTGTATTATTATTTTACTGATGTGTAATAATAATTTGTATAGAATAATGAGAGG 335
|||||

Qy 241 AGAGGA 246
|||||

Db 334 AGAGGA 329
|||||

RESULT 6
A00604

LOCUS A00604 576 bp DNA linear PAT 16-FEB-1993
DEFINITION B.amyloliquefaciens alpha-amylase gene.
ACCESSION A00604
VERSION A00604.1 GI:14540
KEYWORDS alpha-amylase.
SOURCE Bacillus amyloliquefaciens.
ORGANISM Bacillus amyloliquefaciens
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

FEATURES
source Location/Qualifiers
1..576
/organism="Bacillus amyloliquefaciens"
/db_xref="taxon:1390"

BASE COUNT 190 a 98 c 139 g 149 t

ORIGIN
Query Match 97.5%; Score 242.8; DB 6; Length 576;
Best Local Similarity 99.2%; Pred. No. 4.6e-48;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCGCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCCTTTGATGACTGATGATTGG 60
|||||
Db 3 GCGCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCCTTTGATGACTGATGATTGG 62
|||||

Qy 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAGAAAGAACCATATAAAATACCTTGCTCTGC 120
|||||
Db 63 CTGAAGAAGTGGATCGATTGTTTGAGAAAGAAAGAACCATATAAAATACCTTGCTCTGC 122
|||||

Qy 121 ATCAGACAGGGTATTTTATCTGCTCCAGACTGTCGCGTGTAAATAATAGGAATAAA 180
|||||
Db 123 ATCAGACAGGGTATTTTATCTGCTCCAGACTGTCGCGTGTAAATAATAGGAATAAA 182
|||||

Qy 181 GGGGGTGTATTATTATTTTACTGATGTGTAATAATAATTTGTATAGAATAATGAGAGG 240
|||||
Db 183 GGGGGTGTATTATTATTTTACTGATGTGTAATAATAATTTGTATAGAATAATGAGAGG 242
|||||

Qy 241 AGAGGA 246
|||||

Db 243 AGAGGA 248
|||||

RESULT 5
A00156/c

LOCUS A00156 576 bp DNA linear PAT 28-JAN-1993
DEFINITION Nucleotide sequence 2 from patent number GB2133408.
ACCESSION A00156
VERSION A00156.1 GI:14424
KEYWORDS
SOURCE Bacillus sp.
ORGANISM Bacillus sp.
REFERENCE 1 (bases 1 to 576)
AUTHORS Palva,I.

```
Db 243 AGAGGA 248

RESULT 7
LOCUS A00605 576 bp DNA linear PAT 13-APR-1993
DEFINITION B.amyloliquefaciens alpha-amylase gene, reverse complement.
ACCESSION A00605
VERSION A00605.1 GI:14541
KEYWORDS alpha-amylase.
SOURCE Bacillus amyloliquefaciens.
ORGANISM Bacillus amyloliquefaciens
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

FEATURES
    Location/Qualifiers
        source
            1..576
            /organism="Bacillus amyloliquefaciens"
            /db_xref="taxon:1390"
            /db_xref="taxon:1390"
        BASE COUNT 149 a 139 c 98 g 190 t
        ORIGIN
            Query Match 97.5%; Score 242.8; DB 6; Length 576;
            Best Local Similarity 99.2%; Pred. No. 4.6e-48;
            Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTGG 60
|||||
Db 574 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTGG 515
|||||

QY 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAGAGAACACCATAAAATACCTTGTCTGTC 120
|||||
Db 514 CTGAAGAAGTGGATCGATTGTTTGAGAAAGAGAACACCATAAAATACCTTGTCTGTC 455
|||||

QY 121 ATCAGACAGGATATTTTATGCTGTCAGACTGTCGCTGCTGTATTAAGAAATAGCAATAA 180
|||||
Db 454 ATCAGACAGGATATTTTATGCTGTCAGACTGTCGCTGCTGTATTAAGAAATAGCAATAA 395
|||||

QY 181 GGGGGTGTATTATTTACTGATATGTAATAATATATTGTTATAGAAAATGAGAGGG 240
|||||
Db 394 GGGGGTGTATTATTTACTGATATGTAATAATATATTGTTATAGAAAATGAGAGGG 335
|||||

QY 241 AGAGGA 246
|||||
Db 334 AGAGGA 329

RESULT 8
BACAA
LOCUS BACAA 2084 bp mRNA linear BCT 26-APR-1993
DEFINITION Bacillus amyloliquefaciens alpha-amylase gene, complete cds.
ACCESSION J01542 J01543 M12033 M12034
VERSION J01542.1 GI:142428
KEYWORDS alpha-amylase; amylase.
SOURCE Bacillus amyloliquefaciens Marburg strain-ih, cDNA to mRNA.
ORGANISM Bacillus amyloliquefaciens
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
REFERENCE 1 (bases 1 to 537)
AUTHORS Palva, I., Pettersson, R.F., Kalkkinen, N., Lehtovaara, P., Sarvas, M., Soderlund, H., Takkinen, K. and Kaariainen, L.
TITLE Nucleotide sequence of the promoter and NH2-terminal signal peptide region of the alpha-amylase gene from Bacillus amyloliquefaciens
JOURNAL Gen 15 (1), 43-51 (1981)
MEDLINE 82051296
PUBMED 6170539
REFERENCE 2 (bases 76 to 2084)
AUTHORS Takkinen, K., Pettersson, R.F., Kalkkinen, N., Palva, I., Soderlund, H. and Kaariainen, L.
TITLE Amino acid sequence of alpha-amylase from Bacillus amyloliquefaciens deduced from the nucleotide sequence of the cloned gene
JOURNAL J. Biol. Chem. 258 (2), 1007-1013 (1983)
MEDLINE 83108808
PUBMED 6185474
REFERENCE 3 (bases 182 to 237; 1796 to 1857)

Lehtovaara, P., Ulmanen, I. and Palva, I.
In vivo transcription initiation and termination sites of an alpha-amylase gene from Bacillus amyloliquefaciens cloned in Bacillus subtilis
Gene 30 (1-3), 11-16 (1984)
6210229
The deduced amino acid sequence deviates slightly from a published sequence for this species (four conflicts). The signal peptide seems unusually large (31 residues) compared to other known signal peptides. The promoter and mRNA start point are speculatively addressed.
Location/Qualifiers
    source
        1..2084
        /organism="Bacillus amyloliquefaciens"
        /db_xref="taxon:1390"
        /db_xref="taxon:1390"
        /product="alpha-amylase mRNA"
        250..1794
        /note="alpha-amylase protein precursor (EC 3.2.1.1)"
        /codon_start=1
        /transl_table=11
        /protein_id="AA22191.1"
        /db_xref="GI:142429"
        /translation="MIQKRRTVSFRLVLMCTLLFVSLPTTKTSVANGTLMQYFEWYT
        PNDGHWKRLQNDAEHLSDIGITAVWIPAYKGLSDNGYGPYDLYDLGEFOQKGTV
        RTKYGTSELQDAIGSLHSRNVQYGDVYLHNKAGADATEDYTAVEVPAANRQETSE
        EQIKAWTDIFRFGNGTSDPKHWHYHGDADDESRIKSRIFRKGEGKANDWEVS
        SENGNYDLYMADVDYDHPVVAETKKWIGINWLNELSDGFRIDAAKHKFSFLRDMV
        QAVRQATGKEMFTVAEYQWNNAGKLENYLNKTSFNGSFDVPLHFNLAASSGGGYD
        MRLRLDGTVSRHPKAVTFVENHDTQPGQSLSTVQTFKFLAYFILTRESGYDQV
        FYGDMVGTGTSKPEISLKDNIETPLKARKEAYGPOHDYIDHPDVIWGTREDDSA
        AKSGLAALITDPPGSGKRYAGLNAGETWYDITGNRSDTVKIGSDGWEFHYNDGSV
        SIYOK"
        250..342
        /note="alpha-amylase signal peptide"
        343..1791
        /product="alpha-amylase"
        /db_xref="taxon:1390"
        /db_xref="taxon:1390"
    BASE COUNT 610 a 401 c 544 g 529 t
    ORIGIN

Query Match 97.5%; Score 242.8; DB 1; Length 2084;
Best Local Similarity 99.2%; Pred. No. 3.7e-48;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTGG 60
|||||
Db 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTGG 60
|||||

QY 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAGAGAACACCATAAAATACCTTGTCTGTC 120
|||||
Db 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAGAGAACACCATAAAATACCTTGTCTGTC 120
|||||

QY 121 ATCAGACAGGATATTTTATGCTGTCAGACTGTCGCTGCTGTATTAAGAAATAGCAATAA 180
|||||
Db 121 ATCAGACAGGATATTTTATGCTGTCAGACTGTCGCTGCTGTATTAAGAAATAGCAATAA 180
|||||

QY 181 GGGGGTGTATTATTTACTGATATGTAATAATATATTGTTATAGAAAATGAGAGGG 240
|||||
Db 181 GGGGGTGTATTATTTACTGATATGTAATAATATATTGTTATAGAAAATGAGAGGG 240
|||||

QY 241 AGAGGA 246
|||||
Db 241 AGAGGA 246

RESULT 9
LOCUS AR008285 2084 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 3 from patent US 5753460.
ACCESSION AR008285
VERSION AR008285.1 GI:3967394
KEYWORDS
```

```

SOURCE      Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 2084)
AUTHORS      Bisgard-Frantzen,H., Borchert,T.Vedel., Svendsen,A., Thellersen,M.
              and Van der Zee,P.
TITLE        Amylase variants
JOURNAL      Patent: US 5753460-A 3 19-MAY-1998;
FEATURES     Location/Qualifiers
              1. .2084
              /organism="unknown"
BASE COUNT   610 a 401 c 544 g 529 t
ORIGIN

Query Match      97.5%; Score 242.8; DB 6; Length 2084;
Best Local Similarity 99.2%; Pred. No. 3.7e-48;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60

Qy 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120
Db 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180

Qy 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240
Db 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 10
AR037275
LOCUS      AR037275      2084 bp      DNA      linear      PAT 29-SEP-1999
DEFINITION Sequence 3 from patent US 5801043.
ACCESSION  AR037275
VERSION     AR037275.1 GI:5955131
KEYWORDS    .
SOURCE      Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 2084)
AUTHORS      Bisq.ang.rd-Frantzen,H., Borchert,T.Vedel., Svendsen,A.,
              Thellersen,M. and Van der Zee,P.
TITLE        Amylase variants
JOURNAL      Patent: US 5801043-A 3 01-SEP-1998;
FEATURES     Location/Qualifiers
              1. .2084
              /organism="unknown"
BASE COUNT   610 a 401 c 544 g 529 t
ORIGIN

Query Match      97.5%; Score 242.8; DB 6; Length 2084;
Best Local Similarity 99.2%; Pred. No. 3.7e-48;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60

Qy 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120
Db 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180

Qy 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240
Db 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 11
AR052145
LOCUS      AR052145      2084 bp      DNA      linear      PAT 29-SEP-1999
DEFINITION Sequence 3 from patent US 5830837.
ACCESSION  AR052145
VERSION     AR052145.1 GI:5975509
KEYWORDS    .
SOURCE      Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 2084)
AUTHORS      Bisq.ang.rd-Frantzen,H., Borchert,T.Vedel., Svendsen,A.,
              Thellersen,M. and Van der Zee,P.
TITLE        Amylase variants
JOURNAL      Patent: US 5830837-A 3 03-NOV-1998;
FEATURES     Location/Qualifiers
              1. .2084
              /organism="unknown"
BASE COUNT   610 a 401 c 544 g 529 t
ORIGIN

Query Match      97.5%; Score 242.8; DB 6; Length 2084;
Best Local Similarity 99.2%; Pred. No. 3.7e-48;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60

Qy 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120
Db 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180

Qy 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240
Db 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 12
AR087551
LOCUS      AR087551      2084 bp      DNA      linear      PAT 07-SEP-2000
DEFINITION Sequence 3 from patent US 5989169.
ACCESSION  AR087551
VERSION     AR087551.1 GI:10014314
KEYWORDS    .
SOURCE      Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 2084)
AUTHORS      Svendsen,A., Bisq.ang.rd-Frantzen,H. and Borchert,T.Vedel.
              .alpha.-amylase mutants
JOURNAL      Patent: US 5989169-A 3 23-NOV-1999;
FEATURES     Location/Qualifiers
              1. .2084
              /organism="unknown"
BASE COUNT   610 a 401 c 544 g 529 t
ORIGIN

Query Match      97.5%; Score 242.8; DB 6; Length 2084;
Best Local Similarity 99.2%; Pred. No. 3.7e-48;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60

Qy 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120
Db 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180
```

```

Db 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180

Qy 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240
Db 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 11
AR052145
LOCUS      AR052145      2084 bp      DNA      linear      PAT 29-SEP-1999
DEFINITION Sequence 3 from patent US 5830837.
ACCESSION  AR052145
VERSION     AR052145.1 GI:5975509
KEYWORDS    .
SOURCE      Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 2084)
AUTHORS      Bisq.ang.rd-Frantzen,H., Borchert,T.Vedel., Svendsen,A.,
              Thellersen,M. and Van der Zee,P.
TITLE        Amylase variants
JOURNAL      Patent: US 5830837-A 3 03-NOV-1998;
FEATURES     Location/Qualifiers
              1. .2084
              /organism="unknown"
BASE COUNT   610 a 401 c 544 g 529 t
ORIGIN

Query Match      97.5%; Score 242.8; DB 6; Length 2084;
Best Local Similarity 99.2%; Pred. No. 3.7e-48;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60

Qy 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120
Db 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180

Qy 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240
Db 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 12
AR087551
LOCUS      AR087551      2084 bp      DNA      linear      PAT 07-SEP-2000
DEFINITION Sequence 3 from patent US 5989169.
ACCESSION  AR087551
VERSION     AR087551.1 GI:10014314
KEYWORDS    .
SOURCE      Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 2084)
AUTHORS      Svendsen,A., Bisq.ang.rd-Frantzen,H. and Borchert,T.Vedel.
              .alpha.-amylase mutants
JOURNAL      Patent: US 5989169-A 3 23-NOV-1999;
FEATURES     Location/Qualifiers
              1. .2084
              /organism="unknown"
BASE COUNT   610 a 401 c 544 g 529 t
ORIGIN

Query Match      97.5%; Score 242.8; DB 6; Length 2084;
Best Local Similarity 99.2%; Pred. No. 3.7e-48;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60

Qy 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120
Db 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180

Qy 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240
Db 181 GGGGGGTTGTTATTATTTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG 240

Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 12
AR087551
LOCUS      AR087551      2084 bp      DNA      linear      PAT 07-SEP-2000
DEFINITION Sequence 3 from patent US 5989169.
ACCESSION  AR087551
VERSION     AR087551.1 GI:10014314
KEYWORDS    .
SOURCE      Unknown.
ORGANISM     Unclassified.
REFERENCE    1 (bases 1 to 2084)
AUTHORS      Svendsen,A., Bisq.ang.rd-Frantzen,H. and Borchert,T.Vedel.
              .alpha.-amylase mutants
JOURNAL      Patent: US 5989169-A 3 23-NOV-1999;
FEATURES     Location/Qualifiers
              1. .2084
              /organism="unknown"
BASE COUNT   610 a 401 c 544 g 529 t
ORIGIN

Query Match      97.5%; Score 242.8; DB 6; Length 2084;
Best Local Similarity 99.2%; Pred. No. 3.7e-48;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCGCTTTGATGACTGATGATTTGG 60

Qy 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120
Db 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATAAAAATACCTTGCTGTC 120

Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCGCTGTGTAAGAAAATAGGAATAAA 180
```

source		1. .2084		/organism="unknown"			
BASE COUNT	610 a	401 c	544 g	529 t			
ORIGIN							
Query Match	97.5%; Score 242.8; DB 6; Length 2084;						
Best Local Similarity	99.2%; Pred. No. 3.7e-48;						
Matches 244; Conservative	0; Mismatches 2; Indels 0; Gaps 0;						
QY	1	GC	CCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG	60			
Db	1	GC	CCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG	60			
QY	61	CT	GAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTC	120			
Db	61	CT	GAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTC	120			
QY	121	AT	CAGACAGGGTATTTTATGCTGCCAGACTGTCGCGTGTGTAATAATAGGAATAAA	180			
Db	121	AT	CAGACAGGGTATTTTATGCTGCCAGACTGTCGCGTGTGTAATAATAGGAATAAA	180			
QY	181	GG	GGGGTGTATTATTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG	240			
Db	181	GG	GGGGTGTATTATTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG	240			
QY	241	AG	AGAGGA 246				
Db	241	AG	AGAGGA 246				
RESULT 13							
AR129916							
LOCUS	AR129916 2084 bp DNA linear PAT 16-MAY-2001						
DEFINITION	Sequence 13 from patent US 6187576.						
ACCESSION	AR129916						
VERSION	AR129916.1 GI:14117813						
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unclassified.						
REFERENCE	1 (bases 1 to 2084)						
AUTHORS	Svendsen,A., Borchert,T.,Vedel. and Bisg.ang.rd-Frantzen,H.						
TITLE	.alpha.-amylase mutants						
JOURNAL	Patent: US 6187576-A 13 13-FEB-2001;						
FEATURES	Location/Qualifiers						
source	1. .2084						
BASE COUNT		610 a	401 c	544 g	529 t		
ORIGIN							
Query Match	97.5%; Score 242.8; DB 6; Length 2084;						
Best Local Similarity	99.2%; Pred. No. 3.7e-48;						
Matches 244; Conservative	0; Mismatches 2; Indels 0; Gaps 0;						
QY	1	GC	CCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG	60			
Db	1	GC	CCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG	60			
QY	61	CT	GAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTC	120			
Db	61	CT	GAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTC	120			
QY	121	AT	CAGACAGGGTATTTTATGCTGCCAGACTGTCGCGTGTGTAATAATAGGAATAAA	180			
Db	121	AT	CAGACAGGGTATTTTATGCTGCCAGACTGTCGCGTGTGTAATAATAGGAATAAA	180			
QY	181	GG	GGGGTGTATTATTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG	240			
Db	181	GG	GGGGTGTATTATTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG	240			
QY	241	AG	AGAGGA 246				
Db	241	AG	AGAGGA 246				
RESULT 14							
AR137905							
LOCUS	AR137905 2084 bp DNA linear PAT 16-JUN-2001						
DEFINITION	Sequence 31 from patent US 6197565.						
ACCESSION	AR137905						
VERSION	AR137905.1 GI:14479414						
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unclassified.						
REFERENCE	1 (bases 1 to 2084)						
AUTHORS	Svendsen,A., Kjaerulff,S., Bisgaard-Frantzen,H. and Andersen,C.						
TITLE	.alpha.-Amylase variants						
JOURNAL	Patent: US 6197565-A 31 06-MAR-2001;						
FEATURES	Location/Qualifiers						
source	1. .2084						
BASE COUNT		610 a	401 c	544 g	529 t		
ORIGIN							
Query Match	97.5%; Score 242.8; DB 6; Length 2084;						
Best Local Similarity	99.2%; Pred. No. 3.7e-48;						
Matches 244; Conservative	0; Mismatches 2; Indels 0; Gaps 0;						
QY	1	GC	CCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG	60			
Db	1	GC	CCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG	60			
QY	61	CT	GAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTC	120			
Db	61	CT	GAGAAGTGCATCGATTGTTTGAGAAAAGAGAGACCATATAAAATACCTTGTCTGTC	120			
QY	121	AT	CAGACAGGGTATTTTATGCTGCCAGACTGTCGCGTGTGTAATAATAGGAATAAA	180			
Db	121	AT	CAGACAGGGTATTTTATGCTGCCAGACTGTCGCGTGTGTAATAATAGGAATAAA	180			
QY	181	GG	GGGGTGTATTATTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG	240			
Db	181	GG	GGGGTGTATTATTTACTGATATGTAATAATAATTTGTATAGAAAATGAGAGGG	240			
QY	241	AG	AGAGGA 246				
Db	241	AG	AGAGGA 246				
RESULT 15							
AR143262							
LOCUS	AR143262 2084 bp DNA linear PAT 08-AUG-2001						
DEFINITION	Sequence 58 from patent US 6204232.						
ACCESSION	AR143262						
VERSION	AR143262.1 GI:15104548						
KEYWORDS	Unknown.						
SOURCE	Unknown.						
ORGANISM	Unclassified.						
REFERENCE	1 (bases 1 to 2084)						
AUTHORS	Borchert,T.,Vedel., Svendsen,A., Andersen,C., Nielsen,B.,						
TITLE	Nissen,T.,Lauesgaard. and Kj.ae buttet.ruiff.Sslashedren.						
JOURNAL	.alpha.-amylase mutants						
FEATURES	Patent: US 6204232-A 58 20-MAR-2001;						
source	Location/Qualifiers						
source	1. .2084						
BASE COUNT		610 a	401 c	544 g	529 t		
ORIGIN							
Query Match	97.5%; Score 242.8; DB 6; Length 2084;						
Best Local Similarity	99.2%; Pred. No. 3.7e-48;						
Matches 244; Conservative	0; Mismatches 2; Indels 0; Gaps 0;						
QY	1	GC	CCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTGG	60			

```
Db 1 GCGCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTTGG 60
Oy 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAAGAGACCATAAAAAATACCTTGCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAAGAGACCATAAAAAATACCTTGCTGTC 120
Oy 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAAAAATAGGAATAAA 180
Oy 181 GGGGGTGTATTATTTTACTGATATGTAAAAATATAATTTGCTATAAGAAAATGAGAGG 240
Db 181 GGGGGTGTATTATTTTACTGATATGTAAAAATATAATTTGCTATAAGAAAATGAGAGG 240
Oy 241 AGAGGA 246
Db 241 AGAGGA 246
```

Search completed: February 21, 2003, 07:40:17
Job time : 1159 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: February 21, 2003, 07:16:56 ; Search time 2353 Seconds
(without alignments)
2660.637 Million cell updates/sec

Title: US-09-936-145-1
Perfect score: 249
Sequence: 1 gccccgacatacgaaga.....aaatgagaggagagatcc 249

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues
Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main.*
1: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	249	100.0	249	35	US-09-936-145-1
2	249	100.0	270	35	US-09-936-145-2
3	246	98.8	2604	40	US-10-146-327-5
4	242.8	97.5	2084	8	US-08-458-387-3
5	242.8	97.5	2084	10	US-08-600-908-3
6	242.8	97.5	2084	10	US-08-683-838-3
7	242.8	97.5	2084	14	US-09-096-087-3
8	242.8	97.5	2084	17	US-09-325-603-3
9	242.8	97.5	2084	17	US-09-327-563-3
10	242.8	97.5	2084	17	US-09-327-563B-3
11	242.8	97.5	2084	18	US-09-441-313-31
12	242.8	97.5	2084	18	US-09-441-313A-31
13	242.8	97.5	2084	21	US-09-545-586-13
14	242.8	97.5	2084	25	US-09-648-826-3
15	242.8	97.5	2084	25	US-09-648-826A-3
16	242.8	97.5	2084	30	US-09-769-864-58
17	242.8	97.5	2084	32	US-09-854-346-9
18	242.8	97.5	2084	34	US-09-918-543-9
19	242.8	97.5	2084	34	US-09-925-576C-9
20	242.8	97.5	2084	41	US-10-184-771-3
21	242.8	97.5	2084	41	US-10-186-042-3

Qy	181	GGGGGTTGTATATTTTACTGATATGTAATAATAATTTGTATAGAAATGAGAGG	240
Db	181	GGGGGTTGTATATTTTACTGATATGTAATAATAATTTGTATAGAAATGAGAGG	240
Qy	241	AGAGGATCC	249
Db	241	AGAGGATCC	249

RESOL I
US-09-936-145-1

```

; PRIOR FILING DATE: 1999-10-06
;
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Microsoft Word 97 SR-2
; SEQ ID NO 2
; LENGTH: 270
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
US-09-936-145-2

```

Query Match	100.0%;	Score 249;	DB 35;	Length 270;
Best Local Similarity	100.0%;	Pred. No. 1.6e-53;		
Matches 249;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GCCTCCACATACGAAAGACTGGCTGAAACAACTTCAGAGCTTTTGATGACTGATGATTGG	60	
Db	1	GCCTCCACATACGAAAGACTGGCTGAAACAACTTCAGAGCTTTTGATGACTGATGATTGG	60	
Qy	61	CTGAAGAAGTGGATCGATTGTTTGAGAAAGAGAGACCATAAAAATACCTTGTCTGTGTC	120	
Db	61	CTGAAGAAGTGGATCGATTGTTTGAGAAAGAGAGAGACCATAAAAATACCTTGTCTGTGTC	120	
Qy	121	ATCAGACAGGGTATTTTTATGCTGCTCCAGACTGCCGTGTGTAATAAATAGGAATAAA	180	
Db	121	ATCAGACAGGGTATTTTTATGCTGCTCCAGACTGCCGTGTGTAATAAATAGGAATAAA	180	
Qy	181	GGGGGTTGTTATTTACTGATATGAAATAATAATTTGTATAGAAAAATGAGAGGG	240	
Db	181	GGGGGTTGTTATTTACTGATATGAAATAATAATTTGTATAGAAAAATGAGAGGG	240	

```

DD      181  CCGGGGGTGTTATTTATTTAC1GATA1GTAAGAAATAATTTGTTATTAAGAAATAGAGAGGG 241
QY      241  AGAGGATCC 249
          |||||
Db      241  AGAGGATCC 249

RESULT 3
US-10-146-327-5
; Sequence 5, Application US/10146327
; GENERAL INFORMATION:
; APPLICANT: Andersen, Carsten
; APPLICANT: Jorgensen, Christel T.
; APPLICANT: Bisgard-Frantzen, Henrik
; APPLICANT: Svendsen, Allan
; APPLICANT: Kiaerulff, Soren

```

```

; TITLE OF INVENTION: Alpha-Amylase Variants
; FILE REFERENCE: 5886.200-US
; CURRENT APPLICATION NUMBER: US/10/146.327
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US/09/537.168
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: PA 1999 00437
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: 60/127,427
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2604
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
US-10-146-327-5

Query Match          98.8%; Score 246; DB 40; Length 2604;
Best Local Similarity 100.0%; Pred. No. 1.8e-52;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTTGG 60
Db 521 GCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTTGG 580
Qy 61 CTGAAGAAGTGCATCGATTGTTTGAAGAAGAAGAACACCATATAAAATACCTTGTCTGTC 120
Db 581 CTGAAGAAGTGCATCGATTGTTTGAAGAAGAAGAACACCATATAAAATACCTTGTCTGTC 640
Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTGTAATAAATAGGAATAAA 180
Db 641 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTGTAATAAATAGGAATAAA 700
Qy 181 GGGGGGTTGTTATTTTACTGATGATGTAAATATATAATTTGTATAAGAAAATGAGAGGG 240
Db 701 GGGGGGTTGTTATTTTACTGATGATGTAAATATATAATTTGTATAAGAAAATGAGAGGG 760
Qy 241 AGAGGA 246
Db 761 AGAGGA 766

RESULT 4
US-08-458-387-3
; Sequence 3, Application US/08458387
; GENERAL INFORMATION:
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; APPLICANT: Svendsen, Allan
; APPLICANT: Thellersen, Marianne
; APPLICANT: Van der Zee, Pia
; TITLE OF INVENTION: AMYLASE VARIANTS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk of North America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458.387
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/343.804
; FILING DATE: 22-NOV-1994

; TITLE OF INVENTION: Alpha-Amylase Variants
; FILE REFERENCE: 5886.200-US
; CURRENT APPLICATION NUMBER: US/10/146.327
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US/09/537.168
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: PA 1999 00437
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: 60/127,427
; PRIOR FILING DATE: 1999-04-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 2604
; TYPE: DNA
; ORGANISM: Bacillus amyloliquefaciens
US-10-146-327-5

Query Match          98.8%; Score 246; DB 40; Length 2604;
Best Local Similarity 100.0%; Pred. No. 1.8e-52;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTTGG 60
Db 521 GCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTTGG 580
Qy 61 CTGAAGAAGTGCATCGATTGTTTGAAGAAGAAGAACACCATATAAAATACCTTGTCTGTC 120
Db 581 CTGAAGAAGTGCATCGATTGTTTGAAGAAGAAGAACACCATATAAAATACCTTGTCTGTC 640
Qy 121 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTGTAATAAATAGGAATAAA 180
Db 641 ATCAGACAGGGTATTTTATGCTCTCCAGACTGTCCTGCTGTGTAATAAATAGGAATAAA 700
Qy 181 GGGGGGTTGTTATTTTACTGATGATGTAAATATATAATTTGTATAAGAAAATGAGAGGG 240
Db 701 GGGGGGTTGTTATTTTACTGATGATGTAAATATATAATTTGTATAAGAAAATGAGAGGG 760
Qy 241 AGAGGA 246
Db 761 AGAGGA 766

RESULT 5
US-08-600-908-3
; Sequence 3, Application US/08600908
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Bisg id-Frantzen, Henrik
; APPLICANT: Borchert, Torben Vedel
; TITLE OF INVENTION: '-Amylase Mutants
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk of North America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600.908
; FILING DATE: 13-FEB-1996
```

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2084 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 250..1791
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 250..342
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 343..1791
US-08-600-908-3

Query Match 97.5%; Score 242.8; DB 10; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.1e-51;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTGG 60
QY 61 CTGAAGAGTGGATGATGTTTATGCTGCTCCAGACTGTCGCTGTGTAAGAAATAGGAATAAA 120
Db 61 CTGAAGAGTGGATGATGTTTATGCTGCTCCAGACTGTCGCTGTGTAAGAAATAGGAATAAA 120
QY 121 ATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCTGTGTAAGAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCTGTGTAAGAAATAGGAATAAA 180
QY 181 GGGGGTGTGTTATTTACTGATATGTAATAATATATTTGTAAGAAATAGGAATAAA 240
Db 181 GGGGGTGTGTTATTTACTGATATGTAATAATATATTTGTAAGAAATAGGAATAAA 240
QY 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 6
US-08-683-838-3
Sequence 3, Application US/08683838
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben
TITLE OF INVENTION: "-Amylase Mutants
CORRESPONDENCE ADDRESSES: 12
ADDRESS: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/683,838
FILING DATE: 18-JUL-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4394.400-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2084 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 250..1791
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 250..342
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 343..1791
US-08-683-838-3

Query Match 97.5%; Score 242.8; DB 10; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.1e-51;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTGAGCCCTTTGATGACTGATGATTGG 60
QY 61 CTGAAGAGTGGATGATGTTTATGCTGCTCCAGACTGTCGCTGTGTAAGAAATAGGAATAAA 120
Db 61 CTGAAGAGTGGATGATGTTTATGCTGCTCCAGACTGTCGCTGTGTAAGAAATAGGAATAAA 120
QY 121 ATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCTGTGTAAGAAATAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGCTCCAGACTGTCGCTGTGTAAGAAATAGGAATAAA 180
QY 181 GGGGGTGTGTTATTTACTGATATGTAATAATATATTTGTAAGAAATAGGAATAAA 240
Db 181 GGGGGTGTGTTATTTACTGATATGTAATAATATATTTGTAAGAAATAGGAATAAA 240
QY 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 7
US-09-096-087-3
Sequence 3, Application US/09096087
GENERAL INFORMATION:
APPLICANT: Bisgaard-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
APPLICANT: Svendsen, Allan
APPLICANT: Thellersen, Marianne
APPLICANT: Van der Zee, Pia
TITLE OF INVENTION: AMYLASE VARIANTS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESS: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/096.087
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/343.804
FILING DATE: 22-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lowrey Dr., Karen A.
REGISTRATION NUMBER: 31,274
REFERENCE/DOCKET NUMBER: 4054.214-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2084 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 250..1794
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 250..342
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 343..1791
US-09-096-087-3

Query Match 97.5%; Score 242.8; DB 14; Length 2084;

Best Local Similarity 99.2%; Pred. No. 1.1e-51;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTTGG 60
Qy 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120
Qy 121 ATCAGACAGGTTATTTTATCTCTCCAGACTGCCGCTGTGTAAGAAATAGGAATAAA 180
Db 121 ATCAGACAGGTTATTTTATCTCTCCAGACTGCCGCTGTGTAAGAAATAGGAATAAA 180
Qy 181 GGGGGTGTATTATTTTACTGATATGTAATAATATAATTTGTATAAGAAATAGAGAGG 240
Db 181 GGGGGTGTATTATTTTACTGATATGTAATAATATAATTTGTATAAGAAATAGAGAGG 240
Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 8

US-09-325-603-3
Sequence 3, Application US/09325603
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: A-Amlyase Mutants
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Novo Nordisk of North America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York

COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/325.603
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/600.908
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2084 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 250..1791
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 250..342
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 343..1791
US-09-325-603-3

Query Match 97.5%; Score 242.8; DB 17; Length 2084;

Best Local Similarity 99.2%; Pred. No. 1.1e-51;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGCACATACGAAAGACTGGCTGAAACATTGAGCCTTTGATGACTGATGATTTGG 60
Qy 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGCATCGATTGTTTGAGAAAGAAAGAACACCATAAAAATACCTTGTCTGTC 120
Qy 121 ATCAGACAGGTTATTTTATCTCTCCAGACTGCCGCTGTGTAAGAAATAGGAATAAA 180
Db 121 ATCAGACAGGTTATTTTATCTCTCCAGACTGCCGCTGTGTAAGAAATAGGAATAAA 180
Qy 181 GGGGGTGTATTATTTTACTGATATGTAATAATATAATTTGTATAAGAAATAGAGAGG 240
Db 181 GGGGGTGTATTATTTTACTGATATGTAATAATATAATTTGTATAAGAAATAGAGAGG 240
Qy 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 9

US-09-327-563-3
Sequence 3, Application US/09327563
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisg rd-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: A-Amlyase Mutants
NUMBER OF SEQUENCES: 13

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Novo Nordisk of North America, Inc.
;; STREET: 405 Lexington Avenue, 64th Floor
;; CITY: New York
;; STATE: New York
;; COUNTRY: United States of America
;; ZIP: 10174-6401
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/327,563
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/08/600,908
;; FILING DATE: 13-FEB-1996
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Green, Reza
;; REGISTRATION NUMBER: 38,475
;; REFERENCE/DOCKET NUMBER: 4394,204-US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-867-0123
;; TELEFAX: 212-878-9655
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2084 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 250..1791
;; FEATURE:
;; NAME/KEY: sig_peptide
;; LOCATION: 250..342
;; FEATURE:
;; NAME/KEY: mat_peptide
;; LOCATION: 343..1791
US-09-327-563B-3

Query Match 97.5%; Score 242.8; DB 17; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.1e-51;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60
|||||
Db 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60

QY 61 CTGAAGAAGTGGATCGATTTGTTGAGAAAAGAGAGAGACCATAAAATACCTTGTCTGTC 120
|||||
Db 61 CTGAAGAAGTGGATCGATTTGTTGAGAAAAGAGAGAGACCATAAAATACCTTGTCTGTC 120

QY 121 ATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAAATAATAGGAATAAA 180
|||||
Db 121 ATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAAATAATAGGAATAAA 180

QY 181 GGGGGTGTATTATTTACTGATATGTAATAATTAATTTGTATAAGAAAATGAGAGGG 240
|||||
Db 181 GGGGGTGTATTATTTACTGATATGTAATAATTAATTTGTATAAGAAAATGAGAGGG 240

QY 241 AGAGGA 246
|||||
Db 241 AGAGGA 246

RESULT 10
US-09-327-563B-3
; Sequence 3, Application US/09327563B
; GENERAL INFORMATION:

;; APPLICANT: Svendsen, Allan
;; APPLICANT: Bisgaard-Frantzen, Henrik
;; APPLICANT: Borchert, Torben Vedel
;; TITLE OF INVENTION: Alpha-Amylase Mutants
;; FILE REFERENCE: 0776/1F216-US2
;; CURRENT APPLICATION NUMBER: US/09/327,563B
;; PRIOR FILING DATE: 1999-06-08
;; PRIOR APPLICATION NUMBER: 08/683,838
;; PRIOR FILING DATE: 1996-07-18
;; NUMBER OF SEQ ID NOS: 16
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 3
;; LENGTH: 2084
;; TYPE: DNA
;; ORGANISM: B. amyloliquefaciens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (250)...(1791)
;; NAME/KEY: mat_peptide
;; LOCATION: (343)...(1791)
;; NAME/KEY: sig_peptide
;; LOCATION: (250)...(342)
US-09-327-563B-3

Query Match 97.5%; Score 242.8; DB 17; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.1e-51;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60
|||||
Db 1 GCCCGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60

QY 61 CTGAAGAAGTGGATCGATTTGTTGAGAAAAGAGAGAGACCATAAAATACCTTGTCTGTC 120
|||||
Db 61 CTGAAGAAGTGGATCGATTTGTTGAGAAAAGAGAGAGACCATAAAATACCTTGTCTGTC 120

QY 121 ATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAAATAATAGGAATAAA 180
|||||
Db 121 ATCAGACAGGGTATTTTATGCTGCCAGACTGTCGCTGTGTAAATAATAGGAATAAA 180

QY 181 GGGGGTGTATTATTTACTGATATGTAATAATTAATTTGTATAAGAAAATGAGAGGG 240
|||||
Db 181 GGGGGTGTATTATTTACTGATATGTAATAATTAATTTGTATAAGAAAATGAGAGGG 240

QY 241 AGAGGA 246
|||||
Db 241 AGAGGA 246

RESULT 11
US-09-441-313-31
; Sequence 31, Application US/09441313
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Kjruliff, S ren
; APPLICANT: Bisgaard-Frantzen, Henrik
; APPLICANT: Andersen, Carsten
; TITLE OF INVENTION: -Amylase Variants
; FILE REFERENCE: 5709.000-US
; CURRENT APPLICATION NUMBER: US/09/441,313
; CURRENT FILING DATE: 1999-11-16
; EARLIER APPLICATION NUMBER: 09/193,068
; EARLIER FILING DATE: 1998-11-16
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 2084
; TYPE: DNA
; ORGANISM: B. amyloliquefaciens
US-09-441-313-31

Query Match 97.5%; Score 242.8; DB 18; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.1e-51;

0

Le

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/327,563
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 4394.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 2084 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 250..1791
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 250..342
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 343..1791
US-09-648-826-3

Query Match 97.5%; Score 242.8; DB 25; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.1e-51;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60
QY 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAGACCACCAATAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAGACCACCAATAAATACCTTGTCTGTC 120
QY 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAATAAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAATAAGGAATAAA 180
QY 181 GGGGGTGTATTATTTACTGATATGTAATAATAAATTTGTATAGAAAATGAGAGGG 240
Db 181 GGGGGTGTATTATTTACTGATATGTAATAATAAATTTGTATAGAAAATGAGAGGG 240
QY 241 AGAGGA 246
Db 241 AGAGGA 246

RESULT 15
US-09-648-826A-3
Sequence 3, Application US/09648826A
GENERAL INFORMATION:
APPLICANT: Svendsen, Allan
APPLICANT: Bisgard-Frantzen, Henrik
APPLICANT: Borchert, Torben Vedel
TITLE OF INVENTION: Alpha-Amylase Mutants
FILE REFERENCE: 0776/1F216-US2
CURRENT APPLICATION NUMBER: US/09/648,826A
PRIOR FILING DATE: 2000-08-25
PRIOR APPLICATION NUMBER: 08/683,838
PRIOR FILING DATE: 1996-07-18
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 2084
TYPE: DNA
ORGANISM: B. amyloliquefaciens

FEATURE:
NAME/KEY: CDS
LOCATION: (250)...(1791)
NAME/KEY: mat_peptide
LOCATION: (343)...(1791)
NAME/KEY: sig_peptide
LOCATION: (250)...(342)
US-09-648-826A-3
Query Match 97.5%; Score 242.8; DB 25; Length 2084;
Best Local Similarity 99.2%; Pred. No. 1.1e-51;
Matches 244; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60
Db 1 GCCCGGCACATACGAAAGACTGGCTGAAACATTTGAGCCTTTGATGACTGATGATTTGG 60
QY 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAGACCACCAATAAATACCTTGTCTGTC 120
Db 61 CTGAAGAAGTGGATCGATTGTTTGAGAAAAGAGAGACCACCAATAAATACCTTGTCTGTC 120
QY 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAATAAGGAATAAA 180
Db 121 ATCAGACAGGGTATTTTATGCTGTCCAGACTGTCCGCTGTGTAAATAAGGAATAAA 180
QY 181 GGGGGTGTATTATTTACTGATATGTAATAATAAATTTGTATAGAAAATGAGAGGG 240
Db 181 GGGGGTGTATTATTTACTGATATGTAATAATAAATTTGTATAGAAAATGAGAGGG 240
QY 241 AGAGGA 246
Db 241 AGAGGA 246

Search completed: February 21, 2003, 08:46:12
Job time : 2358 secs